

Phonetica

INTERNATIONALE ZEITSCHRIFT FÜR PHONETIK
INTERNATIONAL JOURNAL OF PHONETICS
JOURNAL INTERNATIONAL DE PHONÉTIQUE

EDITOR:

E. Zwirner, Münster in Westfalen

COEDITORES:

R. AVANESOV, Moskva
E. BLANQUAERT, Gent
S. K. CHATTERJI, Calcutta
W. DOROSZEWSKI, Warszawa
E. FISCHER-JØRGENSEN, København
D. B. FRY, London
P. L. GARVIN, Washington, D. C.
A. GEMELLI, Milano
H. GLINZ, Zürich
L. HEGEDÜS, Budapest
R. JAKOBSON, Cambridge, Mass.

K. JIMBO, Tokyo
M. JOOS, Madison, Wisc.
B. MALMBERG, Lund
A. MARTINET, Paris
SH. MORAG, Jerusalem
R. OLESCH, Köln
K. L. PIKE, Glendale, Calif.
A. A. REFORMATSKIJ, Moskva
E. W. SELMER, Oslo
A. SOVIJÄRVI, Helsinki
F. TROJAN, Wien

SECRETARIUS:

W. BETHGE, Münster in Westfalen

1958



Vol. 2, No. 3/4

BASEL (Schweiz)

S. KARGER

NEW YORK

INHALT - CONTENTS - SOMMAIRE

A. E. SHARP (LONDON)	Falling-Rising Intonation Patterns in English . . .	127
L. HEGEDŰS (BUDAPEST)	Neuere Untersuchungen über die ungarischen Affrikaten	153
R. W. THOMPSON (HONG-KONG)	The "th" Sounds and Genetics	184
R. HUSSON (PARIS)	Einige Ergebnisse von Untersuchungen der akustischen Struktur der Vokale	193
A. MAACK (BRAUNSCHWEIG)	Regeln der deutschen Silbenmelodie	199
E. FROESCHELS (NEW YORK)	Zur Frage der Gleichheit von Kau- und Sprechbewegungen	219
<i>Sammelreferate - Surveys - Revues générales</i>		
P. DELATTRE (BOULDER, COLO.)	Les indices acoustiques de la parole: Premier rapport (Fin)	226
Libri		251

*Inhalt der nächsten Hefte - Contents of the Following Numbers -
Sommaire des fascicules suivants*

K. L. PIKE (GLENDALE, CALIF.)	Instrumental Collaboration on a Tepehuan (Uto-Aztecan) Pitch Problem
W. STELLER (KIEL)	Generationsunterschiede im Friesischen
N. L. HADDOCK (IBADAN)	The Tones of Bariba
P. IVIĆ (NOVI SAD)	Die Hierarchie der prosodischen Phänomene im serbo-kroatischen Sprachraum
W. S. ALLEN (CAMBRIDGE)	Indo-Aryan
F. HEDBLOM (UPPSALA)	Recording in Dialect Investigation in Sweden
F. GREWEL (AMSTERDAM)	How do Children Acquire the Use of Language?
A. MAACK (BRAUNSCHWEIG)	Über den Lautwandel nach Fónagy
V. TAULI (UPPSALA)	A Chapter of Estonian Quantity
K. FOKKEMA (AMSTELVEEN)	Phonologia frisica
R. LUCHSINGER (ZÜRICH)	Stottern
C. CLARDY (WASHINGTON, D. C.)	Pampango Phonology
H. ROSÉN (JERUSALEM)	Sibilants and Cognate Phones in One Form of Modern Greek
A. J. VAN WINDEKENS (KESSEL-LO-LOUVAIN)	Phonétique des mots grecs d'origine préhellénique

Falling-Rising Intonation Patterns in English

By ALAN E. SHARP, London

Recent contributors to *English Studies* have drawn attention to a problem of English intonation which has long been in need of clarification. Dr. *W. R. Lee*¹ and Dr. *Maria Schubiger*² have given us their views on what they consider to be the various kinds of 'fall-rise intonations' in English. Dr. *Lee* expresses the hope that his contribution, which he modestly describes as 'tentative and marginal only', may help to stimulate observation and discussion. Since I have myself long found the lack of clarity in this field particularly vexatious, I welcome this opportunity to give some semblance of order to my ideas and to present them here for criticism.

The purpose of this paper is twofold: to try to resolve some of the confusion surrounding this particular chapter of English intonation; and, incidentally, to examine some of the implications of my own interpretation for the writing of continuous texts. If justification is needed for the amalgamation of these two aspects of the problem, it is to be found in my experience that deficiencies of analysis have, not unnaturally, had unfortunate repercussions on the presentation of practical material. The unacceptable treatment of the problem here examined has been for me easily the greatest single cause of dissatisfaction not only with existing descriptions of English intonation, but also with much of the reading material so far published, whether accompanied or not by such descriptions. Inasmuch as the search for greater consistency in the writing of continuous texts, however short, may stimulate more careful obser-

¹ Fall-Rise Intonations in English.

² Again: Fall-Rise Intonations in English.

vation, it is my hope that theory and practice may interact in this field to their common advantage.

It must be emphasized at the outset that attention is here focussed on one small corner only of the field of English intonation, without prejudice or deference to the requirements either of other similarly restricted studies or of any thoroughgoing analysis of the whole subject. In view of the long history of confusion in the matter before us, it is felt in this instance to be of greater urgency to try to reduce this confusion in some measure than to concentrate on giving a wider analysis on the assumption, which in this respect at least is obviously unwarranted, that the data which would form the basis of such an analysis are common property. In the interest of the greatest possible lucidity, the discussion will as far as is feasible be narrowed down to the hard core of the problem, and in particular the following restrictions will apply:

- (i) The falling-rising intonation patterns considered are 'final' in the sentence, in that they are not followed by other so-called 'kinetic tones'.
- (ii) Except where otherwise stated, the type of speech in question is 'unemphatic' and 'unemotional'.

In the course of the exposition of the central problem, certain ancillary topics, in themselves both interesting and important, will of necessity be introduced. The discussion of these will be regarded as secondary for our present purpose, and though some at least of them will, it is hoped, be given fuller treatment on another occasion, they will here be strictly subordinated to the main issue.

Some of the special difficulties of the subject are typified by the twin facts that *Lee*, having stated that 'several types of fall-rise are formally distinguishable', goes on to agree with *Schubiger*'s earlier statement that it is sometimes hard to distinguish two of them; and that *Schubiger* herself, in her reply to *Lee*'s article, whilst distinguishing four types of 'fall-rise', can state that Type 3 is often identical in form with 1 and 2. We may summarize the problem before us as being that of deciding how many types of falling-rising intonation we may usefully recognize and what their differentiae are.

The problem becomes more meaningful to those who believe in polysystemic analysis when the question is posed separately for each major sentence-type in turn. That is to say, we shall do better to ask

in the first place how many types of falling-rising intonation may usefully be recognized as applied to, say, statements; then as applied to 'special' or 'general' questions; and so on. We may well find that the results are similar over a wide range of sentence-types: even then we shall not necessarily equate the similar patterns, especially if they are found to be terms in incongruent commutation systems. The guiding spirit of polysystemic analysis, however, is not so much a desire for complexity for its own sake as a dislike of inaccurate oversimplification and specious simplicity. We need feel no disappointment, therefore, if we find that particular intonation patterns are very widely distributed. In any event, for the purposes of textual transcription it may well be found appropriate to use a monosystemic approach. Since, however, it is in any case obviously desirable to remove all inessential sources of possible confusion, we shall concentrate initially on statements.

We may with advantage further restrict the discussion by straightway excluding *Lee's* Type 3 – his 'fall-rise TUNE'. This pattern, which, as *Schubiger* has pointed out, is merely one form of what is best known as 'Tune II', is seldom confused (though see below) with the type of pattern that is the subject of this paper, namely where the fall takes place, or at least begins, within the stressed syllable, hereafter referred to as the 'focal syllable', on which the falling-rising pattern is effectively 'initiated' – that is, if we ignore as irrelevant any syllables preceding the fall. *Schubiger's* Type 4 she herself characterizes as non-final, and we may therefore exclude this also in accordance with restriction (i) above. We are thus left, as far as the current discussion goes, with *Lee's* Types 1 and 2, and *Schubiger's* Types 1, 2 and 3. We shall in fact be essentially concerned in this paper with a distinction of the type made by *Lee* between his Type 1, which he calls 'the fall-rise TONE', and his Type 2, which he calls 'the fall-rise SEQUENCE'. I agree at least to the point of wishing to distinguish a unitary pattern from a binary one. The unitary pattern I shall here refer to as FR (= 'fall-rise', which term I prefer to restrict to this application), the binary as F + R (= 'fall-plus-rise' or 'fall-followed-by-rise').

It is assumed that readers are familiar with the general outline of the falling-rising pattern or patterns, and in particular with the fact that such a pattern may be initiated (in the sense given above) at any point in a sentence up to and including the last syllable. Where the pattern is initiated (and necessarily terminated) on a final,

perhaps solitary, monosyllable, there is no difficulty of interpretation; and this remains in general true where the 'focus' of the pattern is any syllable of a final polysyllabic word. In these cases we are by general consent dealing with the unitary (even if 'spread-over') pattern, and those who extend their use of the notation \sim or its equivalents to non-final syllables here use this notation consistently. Once, however, even a single word follows the 'focal' word and is orthographically separated from it by a space, the pattern is variously described and/or indicated. This in itself is unexceptionable (nay, in my view highly desirable), but it is here that I maintain that inconsistency is to be found. I suggest, in effect, that not infrequently where the description or notation states or implies that we have FR, we have in fact $F + R$, and vice versa. It is this contention that I am concerned to justify in this paper.

In the type of sentence under discussion we have equally reasonable grounds for postulating either FR or $F + R$. Considerations of patterning and the like allow no priority of interpretation to either. The 'occurrence' of $F + R$ is to be expected in view of the unambiguous occurrence on a single 'substrate' (as I shall call what 'remains' of the sentence-form when the intonation pattern has been abstracted) of $F + F$, $R + F$, $F + FR$, etcetera, some at least of which must be analysed as sequences rather than as unitary patterns since they are not found on a monosyllabic structure: whilst the 'spread-over' of FR is already observable in the case of the plurisyllabic one-word sentence, and the existence of pluriverbal 'tails' – that is, the incidence of the last major prominence (see below) on a non-final word – is attested for F and for R equally.

It is obvious that no question of there being confusion or hesitation between FR and $F + R$ would have arisen if there were a really clear-cut phonetic difference between the exponents of the two patterns. One must conclude that outside the small area of universal agreement referred to above the phonic data in themselves supply either no criteria whatever for setting up two distinct patterns or only such differentiae as are difficult to appreciate. In my own view, we are dealing with the second case; that is, FR and $F + R$ have *very similar* exponents.

We may consider the differences between the two patterns under two heads: a) rhythm b) pitch. In both connexions it is important to iterate that we are discussing 'unemphatic, unemotional statements'. The examples given in this part of the paper will probably

be better appreciated at a second reading after what will be said later about the meanings of the two patterns.

In both types of falling-rising pattern the focal syllable may be said to have *major prominence*. This term must remain undefined here, but should not be exclusively associated with 'kinetic tones'. F + R, however, in contrast to FR, is essentially compound, and requires a further major prominence after that of the focal syllable. This necessitates at least one full rhythmical juncture at some point after the fall. The focal syllable cannot, therefore, be followed only by elements that combine with it into a single rhythmical unit. For example, what may be termed the 'enclitic' exponents of the object pronouns appear to unite in this way with a preceding verbal element. Such a sequence – one type has been called a 'phonemic word'³ – cannot constitute the sole effective substrate for F + R. Thus though we can say /ai v ˈsi:n ˈhim/, we cannot, except in illiterate speech, say /ai v ˈsi:n ˈim/: whereas we *can* say /ai v ˈsi:n im/. This is comparatively straightforward: the difficulty evidently begins when once we leave behind the stereotyped list of recognized 'weak forms'. Even in the case of the pronouns, no 'weak form' is recognized for 'it'. In point of fact, however, the *strong* form is rare in my speech, being very often replaced by 'that'. That is to say, though I can quite happily say /ai v ˈsi:n it/, I should in most contexts tend to say /ai v ˈsi:n ˈðæt/ rather than /ai v ˈsi:n ˈit/, whereas the latter would be frequent and normal in the speech of my wife, who is Scots. Again, in my own speech the form /ju:/, for instance, does not occur unstressed in sentence-final position, so that I say /ai v ˈsi:n ju/, not /ai v ˈsi:n ju:/. The latter form, as used by many London speakers, rings oddly in my ears since I tend to associate final /ju:/ with the prominence found, for instance, in /ai v ˈsi:n ˈju:/. This type of difference would be perhaps most clearly appreciated in examples taken from a dialect in which the 'strong' and 'weak' forms of the pronouns were more markedly different than is the case with some of the standard forms. We may compare, for instance, Devonian /ai v ˈsin ən/ (=standard /ai v ˈsi:n im/) with /ai v ˈsin ˈi:/ (=standard /ai v ˈsi:n ˈhim/).

The difficulty is further aggravated when we move from pronouns, which as a class are acknowledged to have 'weak forms', to, say, nouns, which in general are not. The occurrence of rhythmically

³ Trager and Smith, § 1. 73.

weak forms of, e.g., 'John' is not sufficiently recognized. This is because of the comfortable permanence of the transcriptional symbolization of such words combined with the practice of grammatical spacing, the phonetic validity of which tends to be questioned only when we have, as it were, already been jolted out of our complacency by the occurrence of a recognizable 'weak form'. It tends to be assumed, in fact, that nouns will be stressed. Hence, I do not doubt, much of the confusion between /ai v ˈsi:n dʒɒn/ and /ai v ˈsi:n ˈdʒɒn/. In fact, however, there is a rhythmical relationship between these two *comparable* with, but less easily appreciated than, that between /ai v ˈsi:n im/ and /ai v ˈsi:n ˈhim/.

The presence of a second major prominence, though a prerequisite for the occurrence of F + R, does not, unfortunately, identify it. This is because it is possible to have such prominence *within the rise* of FR. Thus, though /ai v ˈsi:n ju:/ does not occur in my speech, what I shall write as /ai v ˈsi:n ,ju:/ does. This has the same rhythmical pattern as /ai v ˈsi:n ˈju:/. In order to make these distinctions more credible, I cannot avoid some passing reference to a subject which will require full treatment of its own elsewhere – prominence in so-called intonation 'tails'. We find not only /ai v ˈsi:n im/ but also /ai v ˈsi:n ,him/, where /, / is used to denote major prominence on a (*in the context of /ˈ / only*) low level pitch. This structure, which I shall call F + L, has the same rhythmical pattern as F + R, as in /ai v ˈsi:n ˈhim/, and is subject to the same restrictions, etcetera. I can, for instance, say /ai v ˈsi:n ,ju:/, though not /ai v ˈsi:n ju:/: but for /ai v ˈsi:n ,it/ (normal for my wife) I usually say /ai v ˈsi:n ,ðæt/ (or /ai v ˈsi:n ˈðæt/). Whereas, however, we have /ai v ˈsi:n im/ corresponding to /ai v ˈsi:n im/, we have apparently only one pattern, which might be written either /ai v ˈsi:n ,him/ or /ai v ˈsi:n ˈhim/, to set against /ai v ˈsi:n ,him/ and /ai v ˈsi:n ˈhim/. In other words, the tonal commutation system after /ˈ/, which it is no part of this paper to describe in detail, is apparently not congruent with that after /ˈ/. The 'indeterminate' structure in question I shall call FR + L, using the notation /ˈ, /, to point the parallelism with F + L, since in both cases the *second* prominence may be regarded as not disturbing the outline of the intonation contour already established.

How, then, are we to distinguish F + R from FR + L, since they have the same rhythmical pattern and are both falling-rising? For the answer, insofar as there *is* an answer, we must turn to considerations of pitch. I am conscious in this matter of the danger of

failing to get rid of that 'personal equation' against which we are warned by *Sweet*⁴. Chastened by this warning, although I am fairly certain that what I am about to describe is not just a characteristic of my personal idiolect, I shall for the moment confine myself to saying that *in my own speech at least* there is a distinct tendency for F + R to fall to a lower pitch, and finish on a lower pitch, than FR. In terms of tonic sol-fa notation a typical comparison might be between m s d for F + R and m l r for FR, and the difference may be greater than this. If the 'register' of the second half is comparatively low, I get a distinct impression of F + R: if comparatively high, of FR. I do not pretend, however, that there is a really *clear-cut* difference of this kind. There seem to be many utterances which fall into an area of indeterminacy where considerations of pitch would allow either interpretation.

Thus far attention has been directed to FR and F + R as applied to very brief substrates only. This restriction has been observed advisedly, since the detailed consideration of long substrates would involve us too deeply in those questions of prominence which, as I have already said, merit a separate paper. It may, however, be said here that in addition to the typically higher register, longer FR sentences have, as *Lee* says, a *tendency* to start rising immediately after the fall, though it cannot be denied that the pitch often remains low level over a number of syllables. In the case of F + R, however, in addition to the 'lower' tendency in general, any syllables intervening between the F and R focal syllables regularly remain low, thus *lower* than the low syllables of FR.

In order to obtain, if possible, a somewhat less subjective impression of the difference between the two patterns, I recorded with as little exaggeration as possible fifty sentences, some with FR and some with F + R. I later played back the recording to see how far I could distinguish what I had recorded. In the case of approximately two thirds of the examples I received a distinct impression one way or the other: in the remaining cases I did not feel sufficiently certain to write anything down. Reference to the text of the recording showed that where I had hazarded a guess I had invariably drawn the right conclusion.

Now it is quite obvious that a speech community could not function as such if, *in general*, a speaker did not supply his listeners

⁴ *Sweet*, Preface, VIII.

with sufficient clues to act upon, whatever the nature of the response, if any, desired. Indeed, so apparently lavish is the supply in some cases that it has been considered worthwhile to calculate the degree of 'redundancy' in various languages. Over considerable stretches of speech, then, ambiguity is unlikely to arise. It is recognized, however, that where syntactical criteria are used to establish elements of linguistic structure, such criteria are, by their very nature, not necessarily available to identify the exponents of those elements in any given utterance, so that homophony cannot invariably be resolved. It is nevertheless all too widely assumed, in my view, that the exponents of given elements of structure either regularly are, or regularly are not, homophonous (in a given context).

Whatever may be the phonological factors into which we analyse the totality of speech, we can set up commutation systems of one sort or another to account for the phenomena we observe. In the case of many very simple structures, such as the one-word sentence, we have normally little difficulty in counting the terms of those systems or in identifying their exponents. In the case of longer and more complex structures, however, and even in that of some that are comparatively simple, it is difficult to decide with any certainty whether or not we are dealing with homophony. It seems to me little more than a subterfuge to pretend that this is a question of style, as with, for instance, *Hockett's* antithesis of 'clarity norm' and 'frequency norm'⁵, unless we set up an unusually artificial style from which all phonological ambiguity is by definition excluded. Just as there are cases where the hearer is, so to speak, overwhelmed with multiple clues, so there are cases where despite careful speech and careful listening the available clues prove inadequate. Where failure to distinguish exceeds what is attributable to occasional inadvertence, yet some utterances are unhesitatingly and correctly assigned to one or other rival interpretation, we are in the realm of what may be called (since the term 'overlapping' has been used in another sense) 'shading' exponents.

Many possible cases of such shading suggest themselves in addition to the one before us. There is a whole twilight world here, all too familiar to the punster. I may perhaps refer to the experiment in stress perception made by *N. C. Scott*⁶ which showed that /'impɔ:ts/ and /im'pɔ:ts/ were distinguished only with difficulty on a low level

⁵ *Hockett*, p. 220.

⁶ *Scott*, p. 44.

pitch. A more satisfactory example from my own speech would be the unusual type of distinction operating within a single grammatical category between /'lukaut/ (man on watch) and /luk'aut/ (prospect) as in "That's your own look-out" said with falling pitch on 'own' and low level pitch on 'look-out'.

This type of partial indeterminacy need cause the linguist no serious embarrassment except in the early stages of trying to devise a minimum graphetic vocabulary. Once it has been determined by some rigorous technique what utterances are definitely *not* descriptively equivalent, and such a minimum vocabulary has been chosen to cover all such cases, there will comparatively seldom be hesitation even in recording an individual utterance, since we are in no small measure conditioned in our listening by considerations of expectancy. In any case, where we are concerned, as in transcription from orthography or in the provision of reading material, with supplying the means for adequate renewal of connexion not with some individual utterance but with a whole class of descriptively equivalent utterances, we shall have fulfilled that requirement once we have stated the structure of that utterance-form together with all necessary reading conventions: but these conventions shall include, if it is desired that repeated renewals of connexion shall simulate the variety of the original data, a statement of the permissible degree of overlap, if any, of the exponents of such diversely symbolized elements as the structure may contain. This amounts in practice to saying that where two utterance-forms are distinguishable by ear only part of the time, they may nevertheless be consistently distinguished in transcription.

This assumes, however, the ability to 'recognize' the utterance-form itself: and in a field so ill-charted as intonation it is perhaps not surprising that this ability is often suspect. In an attempt, therefore, to reduce such indecision as may apply to the type of material considered in this paper, an appeal will now be made to those features of a wider context that have been variously taken to constitute 'meaning' of one sort or another.

Such an appeal does not necessarily entail the abandonment of further formal analysis and an immediate recourse to vague notional criteria. Insofar as it is possible to make any clear-cut dichotomy between the intra-linguistic and extra-linguistic components of meaning, it is a widely accepted principle of modern linguistics that reference to the former shall have priority over reference to the latter.

This means for our purpose that the ultimate justification for stating that two intonation patterns differ in meaning will preferably lie in the demonstration by a technique of 'longer piece' analysis that the sentences which they respectively characterize stand in different formal relationships to surrounding speech events; and the statement of their meanings will consist precisely of the description of those relationships. This would be in part no more than a kind of international syntax, but would also involve, for instance, reference to the repetition or non-repetition of particular lexical items. If it ever becomes possible adequately to categorize the non-verbal behaviour associated with speech, further, quasi-formal, correlations may eventually be established. Meanwhile, long acquaintance with such recurrent patterns of verbal and non-verbal behaviour makes, or should make, a sensitive native speaker intuitively aware of the broad syntactical, collocational and situational potentialities of a given utterance-form. In the present unsatisfactory situation it would be doctrinaire to refuse to make use of such resources. As a shorthand substitute for the type of description ultimately desirable it has been found useful in the past to make use of intralingual translation and/or brief indication of actual or potential verbal contexts, and in what follows I shall not hesitate to make use of this method where convenient.

At this point, then, I shall assume that a formal distinction has been established between FR and $F + R$, and shall pass on to consider what generalized meanings if any may be said to characterize the two patterns. Failure to observe a regular distinction between FR and $F + R$ has led to confusion of the meanings properly belonging to the separate patterns and consequent inability to perceive such consistency as is observable in their use. Even now we shall not necessarily have eliminated the problem of homophony, since even if we concede the distinction between FR and $F + R$, we may still feel it necessary to say that we have more than one FR and/or more than one $F + R$ if the meanings covered by either appear to cover too wide a range or be wholly 'discontinuous'. Nevertheless it should now be possible to subdivide the meanings found for all falling-rising sentences into two sections in such a way that any given meaning falls unambiguously into one or other section. The division will be made on the basis of the formal characteristics outlined above. Though in listening to any given utterance we may not apprehend the special characteristics of either pattern, we shall yet find them

over a large enough sampling of such utterances; and having once found them and allotted the meaning accordingly we shall be able to allot any ambiguous utterance we hear to one or other pattern by using the meaning as a clue. If we ourselves are writing texts for others to read, we must choose accordingly, since to anyone who associates different meanings with the two patterns the indication of the wrong pattern will make as little sense as if he were asked to read "Drink up your bier".

The term 'contrastive' might with reason be applied to any element of speech inasmuch as its very occurrence in an utterance contrasts with the occurrence of any other element that might have filled the same place in structure. Consequently any sentence might be held to indicate at least as many contrasts as it contained elements of structure. Nevertheless, those elements in a sentence which may be considered to be in any respect more prominent than their neighbours, though they are not in consequence any *more* contrastive than those neighbours, may perhaps be held to be singled out for special mention and a *highlighting* of such contrasts. Falling-rising intonation patterns have very generally been credited with favouring such highlighting to a particularly high degree – with being, so to speak, the contrastive intonation par excellence. I feel that in the case of *FR* this view is not very wide of the mark. Certainly one very common meaning of *FR* may be described as contrast of the 'this, not that', 'this, but not that' or 'this, even if not that' variety. Thus, for example, "I've 'seen him" (but I've never heard him speak); "'Friday" (not Thursday, as you have just incorrectly stated); "He's 'clever" (even if we don't like him). The contrast, it should be noted, need not be felt to apply only to the word containing the focal syllable. Thus "I've peeled the po'tatoes" may as often suggest "but I've not made the beds" as "but I've not peeled the apples".

The kind of demonstration that would help to put this meaning of contrast on a more acceptable formal basis would be to show that *FR* frequently occurred followed by 'but', 'although', 'even if', and so on, or in contexts where the next clause was a repetition of the first apart from some one word or phrase and the presence or absence of negation, as in *Palmer's* example, quoted by *Lee*, "I didn't say it was 'white, I said it was 'black". It is hoped that this paper may prove to be an earnest of such research.

The general meaning of 'contrast' may probably be stretched to embrace most of the occurrences of *FR*. Even *Lee's* "Sir!" (if we

may temporarily go outside the field of statements), in which he says there is nothing contrastive, might be said to have a flavour of "It's you, not someone else, that I'm calling".

It is less easy to give a concise label to the meaning of $F + R$. *Schubiger* is perfectly right when she speaks of the rise after the fall as giving a certain amount of prominence to the word on which it falls. This might, unfortunately, be said of *any* stressed syllable. Elsewhere⁷ *Schubiger* states that "there is a connection between the pattern of English intonation and the established word-order. It is the scheme of intonation that has brought about in the past the present English word-order so far as it is fixed, and it is this same scheme of intonation that influences word-order today in those cases where it is still free." Whether we can assert this priority of causation is open to question, but we may willingly accept that word-order and intonation are in some way connected. The examples most frequently given for $F + R$ are those where the rising element is relatively independent and may with much the same effect be found before the falling element or separated from it by a pause. This is particularly the case for adverbial words, phrases or clauses. Where, on the other hand, the order of the elements of the substrate is relatively fixed, $F + R$ is not recognized with anything approaching the same consistency. Yet $F + R$ occurs with substrates of many different kinds, and the independence of the rising element from the point of view of word-order is variable. One important example must suffice. In sentences containing (to use the traditional terminology) subject, verb and object it is unusual, though not impossible, for the object to stand first. Thus "John I've seen" and "London I dislike" occur much less frequently than "I've seen John" and "I dislike London": whereas, for instance, "Sometimes I walk" and "Unfortunately, he drinks" compare much more favourably with "I walk sometimes" and "He drinks, unfortunately".

Now with the simple falling pattern ("Tune I" of *Armstrong* and *Ward*) the fall may take place at any point of the substrate in accordance with the requirements of the situation or of idiom, and this may leave any number of words after the fall on a low level pitch. In general these words will be at least partially "anaphoric" in effect, and might in *some* cases be omitted without much loss. In many cases, however, the words after the fall are not so 'superfluous',

⁷ *Schubiger*, 1935, p. 11.

and require a degree of highlighting which cannot be attained on a low level pitch, and which for certain substrates may be 'brought about'⁸ by front-shifting of the final element. Where this front-shifting is not possible, or even where it is, the highlighting may 'alternatively' be brought about by a rise on the final element in question. It must be clearly understood that I do not say that the two word-orders are identical in meaning. From one point of view the very fact that there may be the two possibilities might itself be taken to establish two *different* meanings. My use, therefore, of "with much the same effect" and "alternatively" must be taken to imply nothing more than a close similarity of meaning. Now the relationship of $F + FR$ to $FR + F$ is probably more readily appreciated than that of $F + R$ to $R + F$, and it may be helpful to set up the following equation:

$$F + FR : FR + F :: F + R : R + F$$

or /ai v`si:n`dʒɒn/ : /`dʒɒn ai v`si:n/ ::
/ai v`si:n`dʒɒn/ : /`dʒɒn ai v`si:n/

I should not like to press this equation unduly at the moment, and I do not wish to lay myself open to a charge of *ignotum per ignotius*. I may, however, point out here that front-shifting of this kind with $F + R$ achieves the same result as *Schubiger's* 'stopping in the middle': that is, we find out whether in the new circumstances the focal syllable has a fall only or a fall-rise. I have myself long recommended these tests to students, though I fear that their utility may be very limited, since it may well be that they can be correctly carried out only by those who already appreciate the distinction and have in consequence no need of them. It may nonetheless be worthwhile to add that an expansion of the front-shifted element may prove helpful to some. This expansion may often be given the form "Where'... is/are concerned,...". Thus "I`like 'John" may be referred to "Where 'John is concerned, I`like". Other quasi-expansions might be "My feeling towards 'John is one of `liking" or, if the original, rather than the front-shifted order is preferred, "It is`liking that I feel towards 'John".

In common with simple F , $F + R$ may *sometimes* suggest the addition to the substrate of such words as 'even', 'also' or 'for example'. In such cases the element pinpointed by the fall is asso-

⁸ The use of the expressions "brought about" and "front-shifting" is for convenience of statement only, and must not be taken to imply the primacy of either word-order.

ciated in some way with other such elements. This is in sharp contrast to the often strongly *dissociative* effect of FR. We may compare /^ˆthat's not 'bad/ (something else was quite good: this too is very fair) with /[˘]that's not bad/ (though the others were poor: or, we take no responsibility for the others).

What has been said about word-order, front-shifting and the like may perhaps help to make acceptable the last, and in many ways the most important, point that I should like to make about F + R. It is sometimes said that falling tunes have finality, whereas rising tunes are suspensive. The validity of these statements will neither be upheld nor be challenged. What I should like to emphasize is that F + R has every bit as much finality as any simple F sentence, and nothing of any suspense that may be found in simple R sentences. In other words, *there is no 'implication'*. If F + R *must* be associated with Tune I or Tune II and if the generalizations about these patterns just referred to have any value whatever, then it is emphatically with the *former* that it has the greater affinity.

Further exemplification of the two patterns, together with appropriate comment, will, I hope, help to clarify the distinction that I have been trying to make. The sometimes compressed, sometimes more long-winded explanations in parentheses are themselves exemplificatory only, and must not be taken to exclude other possible meanings equally in contrast with any suggested for the other pattern. Here as elsewhere in this paper all stress and tone marks not immediately relevant have been kept to a minimum to avoid involvement in those questions of prominence to which reference has already been made.

For ease of future reference, my own examples will from this point onwards be continuously numbered. In accordance with what was said earlier, the first examples will all be 'statements': those that follow will be grouped into sections corresponding to other broad categories of sentence-form. This grouping is naturally arbitrary: it is obvious that the sentences of English could be almost endlessly subdivided. The best method of subdivision can, however, be determined only on the completion of the detailed study which I hope may one day be undertaken. The rough labels given to the interim classification are, not, however, notionally based: they summarize formal marks of broad categories of sentence structure which need not be re-stated here. The circularity inherent for our purpose in the use of intonational criteria is necessarily avoided.

Statements

1. I ˘thought ˈso (you have confirmed what I thought).
2. I ˘thought so [(i) but have since been proved wrong;
(ii) but your words have made me doubtful].
3. I ˘think so (but may be wrong).
It is interesting to note that I ˘think ˈso is a most unlikely utterance, and would have to mean something like, "Such is the way my ratiocinatory processes operate".
4. He ˘wants to (but can't).
N.B. He ˘wants ˈto is impossible.
5. He ˘wants ˈtwo (so he can't throw them *all* away).
6. He ˘wants two (but can't afford them).
7. He ˘shot himˈself [(i) his wife he poisoned;
(ii) his own hobby was shooting].
Note that in case (ii) a comma would be usual in the orthography: but neither a pause nor the retention of the /h/ is necessary in the pronunciation. Similarly, "That's the idea, is it?" may be heard as /ðætʰs ði aɪˈdiə ˈiz ɪt/, that is, with the 'intrusive' /r/.
8. He ˘shot himself [(i) but did not die of the wound;
(ii) you're wrong in saying that he cut his wrists].
9. ˘I can do ˈthat (so there's no need for you to boast).
10. ˘I can (you have just said that *you* can't).
11. ˘I ˈcan (unenthusiastic admission of ability in response to, for instance, "Who plays bridge?").
12. ˘He's a ˈplumber (ask *him* how to mend the tap).
13. ˘He's a plumber (the other man isn't).
14. He's a ˘queer ˈchap (as chaps go, he's queer).
15. He's a ˘queer chap [(i) So look out!;
(ii) but really very nice].
16. She gave me the ˘sweetest ˈsmile. (Some sort of smile was, perhaps, to be expected: hers was exceptionally sweet.) Cf. *She gave me the ˈsweetest ˘smile.* (Her smile was indeed very sweet: but, more important, she *might* have scowled.)

Both our description of the two patterns and our examples have so far related to a single type of sentence. A similar distinction may, however, usefully be made in the case of other sentence-types; that is, we can distinguish two falling-rising patterns the characteristics

of which are in general similar to those of FR and F + R as set up for statements. Since it is essentially the *distinction* (or distinctions) that we are concerned with in this paper, I shall keep the same labels in discussing the other sentence-types. Some of the latter, however, tend to involve emotional or emphatic utterance more frequently than statements, so that what has been said above in connexion with the pitch and rhythm characteristics of unemotional, unemphatic statements will in any case not apply with the same regularity here. The distinction itself, however, is no less important; and insofar as it is possible to abstract those features of pitch range, tempo, and so on that correlate with emotional or emphatic utterance, the essential relationship still seems to hold.

Requests

17. 'Eat well (warning: or you'll soon be exhausted).
18. 'Eat 'well (entreaty: there's a good boy!).
19. 'Do'smoke (make yourself at home. I know you like smoking).
20. 'Do smoke (e.g. A to B: 'Don't 'smoke. C to B: 'Do smoke, i.e. ignore A's plea).
21. 'Be an'angel (please).

Thus with this type of sentence it is F + R that gives the effect of entreaty. FR often implies warning, but may occasionally, as with FR elsewhere, be used in correction.

Questions

There seems to be something of a conspiracy of silence where the association of falling-rising patterns with questions is concerned. It is decidedly singular that *Palmer*, whose work on English intonation is not to be lightly dismissed, should have stated⁹ that the fall-rise was never used for questions. *Lee* states¹⁰ that 'questions cannot, seemingly, end with a fall-rise *tone*'; and *Schubiger*¹¹ that 'the occasional fall-rise intonation of questions constitutes too subtle a problem to be treated in this short article'. It would seem, therefore, that this is the most difficult part of our problem, and I must admit that I am not yet certain how far one might challenge

⁹ *Palmer*, p. 82.

¹⁰ *Lee*, p. 70.

¹¹ *Schubiger*, 1956, p. 159, n. 7.

Schubiger's use of 'occasional'. That is to say, I am not yet prepared to say *how often* FR and F + R are used in questions: but that they *are* used is beyond doubt. Whether, again, FR is particularly rare as applied to the final word of a question, I cannot at present say: but utterances of this kind do occur. What is perfectly clear is that in both 'yes-no' questions and 'special' questions at least one focus for the patterns is quite common: the *first* word. FR, in these circumstances, asks for confirmation or repetition: F + R pleads for an answer (or for action). Other possible uses of F + R suggest themselves, but further data will have to be collected before they can be established as genuine. In view of the doubt cast on the use of these patterns for questions, I shall limit my examples here to utterances that I have personally heard and noted down.

a) 'Yes-No' Questions

22. Are 'all your family clever? (*You* obviously are).
23. Have we 'got any coffee? (It's all very well talking about making coffee, but...).
24. Have you 'much luggage? (A little can easily be conveyed).
25. 'Are you 'coming? (Despair: Must I wait here for ever?).
26. 'Will you sit 'down? (Pleading or impatient according to voice quality, tempo, etc.).

b) 'Special' Questions

27. 'What shall I 'tell him? (I really cannot think of anything).
28. 'How do you like your 'coffee? (Intimacy).
29. 'What's his name? [(i) I have forgotten;
(ii) I am incredulous].
30. When 'is he coming? (Several possible dates have already been rejected).
31. How's the 'family? (The remainder of this conversation was not heard, but may have included a previous enquiry about the health of the person addressed).

The use of F + R in the type of pleading question just exemplified seems tantamount to prefixing to the question the words 'Do tell me... Thus 'Are you 'coming? seems equivalent to 'Do tell me whether you're 'coming, with the entreaty superimposed, as it were, on the question. Further elaboration of the meanings of the two

patterns as applied to requests and questions will not be undertaken: but it may be said that the 'contrastive' meaning of FR seems no less evident under these headings, and the reversal and expansion techniques may be applied to at least some of the F + R examples with results not essentially different from those obtained above. Thus even *'Do 'smoke* may be turned into *As far as 'smoking goes, 'do so*; and it is worthy of note that *'do* is sometimes added as an appendage to requests, thus: *Sit 'down, 'do*. It seems at least quite clear that when used as entreaties such sentences are *not* cases of FR.

As an indication of just how far this distinction may be carried for purposes of textual transcription, I append a few miscellaneous examples.

32. Thank you 'so 'much.
33. 'So much? (Is 'this the quantity you require?).
34. He went 'home, of course (but that's as far as he got).
35. He went 'home, of 'course (how stupid of you to ask me where he went!).
36. 'Good old 'Joe!
37. 'All 'right (Resignation).
38. You've 'seen him, haven't you? (Reassure me).
39. You've 'seen him, 'haven't you? (Confirm).
40. Hm! A 'gloomy place! (Whatever its virtues may be).
41. Hm! A 'gloomy 'place! (As places'go, its'gloomy).

We have seen that FR and F + R may usefully be set up for several broad categories of substrate. Are there, it may be asked, any noteworthy grammatical (as opposed to lexically determined) limitations on their incidence? This is a question that requires detailed study, but a few preliminary remarks may not be out of place.

Since F + R is essentially binary, it is not to be expected that it will occur with substrates that are essentially unitary, such as single words. Even compound nouns fall, in general at least, under the same ban: where certain types might be thought to occur with F + R, their status as single words becomes suspect and we get, as it were, the impression that the hyphen has dissolved. Thus one might, for instance, have *It was an 'ordinary 'knife, but a 'fish (-) 'fork*. *Blackbird* quite clearly could not stand in such a sentence

without losing its identity completely. Thus occurrence or non-occurrence with F + R – or, indeed, any other compound pattern – may well be relevant to the *definition* of compound nouns.

Again, we have already seen that a combination of verb and clitic pronoun cannot have F + R, and the same restriction applies where the element on which the rise would fall is a ‘prop-word’ of some kind which cannot normally be stressed. Thus though it is possible to say *I’ve got one*, *one* has here its full numerical status.

The incidence of F + R on those verb and ‘preposition’ sequences that are often analysed as forming together a single constituent has the same disruptive effect as with the compound nouns mentioned above, and necessitates a different analysis. Thus we may say *I’ll eat with you* (but will not accompany you to the theatre afterwards) or (rarely) *I’ll eat with you* (our shared activity shall be eating), but though *I’ll bear with you* may mean that I’ll try to be tolerant, *I’ll bear with you* suggests communal portage.

FR, on the other hand, covers a very wide range, and no limitations come to mind that seem to be absolute. Such partial restrictions as suggest themselves seem to relate to the feasibility of association with elements that are in some way ‘incompatible’ with the ‘contrastive’ function of FR.

In conclusion, I shall give a brief commentary on the treatment of this question by previous writers, ending with the articles by *Lee* and *Schubiger* mentioned at the beginning of this paper. I hope that this commentary, by showing where I feel at variance with previously expressed opinions and the examples that illustrate them, may serve to make clearer my own point of view.

*Kingdon*¹²

Kingdon makes, in effect, no distinction between FR and F + R. It might seem at first sight as if his ‘undivided’ variety of Tone III were my FR and his ‘divided’ and ‘separated’ varieties were my F + R. Indeed, his ‘undivided’ examples are all acceptable as FR, and his ‘divided’ and ‘separated’ examples seem to be in general F + R. This equation will not stand, however, because a) his ‘undivided’ examples are all fairly short, and ‘expanding’ them appears to involve a change to the ‘separated’ variety; (b) the

¹² *Kingdon*, 1948.

'divided' and 'separated' examples are not all acceptable as F + R (e.g. on page 120 *'Somebody must have 'taken it, which I cannot contextualize except as FR – 'Somebody, etc.); c) it is clearly stated that "the feeling and effect of the overall tone remains... the same in all varieties; the unity of the tone is kept intact"*.

Elsewhere¹³, with reference to *There was the 'greatest 'sarcasm in what 'John said*, which appears from the context to be F + R, Kingdon says, "the low rise on John converts the tune into a Tune III which invites the listener to go on thinking for himself on the lines suggested by the statement, though in this case no particular implication is involved". This shows once again the confusion of meanings which bedevils the whole subject. Kingdon cannot, it would seem, put his finger on any particular implication, but feels there ought to be one.

*Armstrong and Ward*¹⁴

In *Armstrong and Ward* the reduction to two tunes is carried to the point where the term 'fall-rise' creeps in, as it were, only by accident. The examples shown with falling-rising patterns in the Handbook of English Intonation are divided between those given under the heading "Tune I followed by Tune II" and those given under the heading "Variation of Tune II, due to special prominence being given to one or more ideas". The former (p. 35) are all cases of final adverbial expressions of one kind or another, which, as has already been said, are more regularly diagnosed as F + R than any other type of sentence. They are said to "share the characteristics of Tune II – i.e. lack of definiteness, a possibility of adding further remarks". The inadequacy of this type of description has already been pointed out. The second group includes statements, requests and questions, the first two being shown with medial syllables, if any, on a low level pitch, the questions with a rise beginning immediately after the fall. The requests are all acceptable as F + R and the questions as FR, but the statements (pp. 58–61) are not homogeneous. For instance, the first two examples on p. 59 – *'Talking wouldn't have been any 'good* and *I 'can't make it without 'eggs* – are almost certainly intended as FR and F + R respectively, whereas the third – *She 'said it didn't 'matter* – could equally well be taken

¹³ Kingdon, 1949, p. 181.

¹⁴ A Handbook of English Intonation.

as either FR (but she was really very cross about it) or F + R (so why didn't you carry on?). This 'variation of Tune II' is stated to show "in all cases a lack of finality". For such of the examples at any rate as are F + R this cannot be maintained.

*Jones*¹⁵

Jones feels the usual difficulty. In § 1056 of the 'Outline of English Phonetics' we read: "The intonation described in §§ 1051-55 (a modification of Tune II, according to *Jones* - AES) is sometimes used in cases where there does not appear to be any obvious contrast, and where it is difficult to specify the reason for the use of the intonation. Notable cases are expressions of regret and entreaties or urgent requests." Precisely: it is difficult to specify the reason for the use of the intonation, because that intonation (FR by implication) is, in fact, *not* used.

*Jassem*¹⁶

Despite his elaborate subdivision of English intonation patterns, *Jassem's* falling-rising examples are as unsatisfactory as any yet published. A detailed examination of those examples would be out of place here: but it is worthy of note that in connexion with his 'low falling NT (= nuclear tone - AES) followed by the low rising NT' (p. 81), of which the instances given are very mixed, he himself says, "*Palmer's* Rising-Falling-Rising Nucleus seems to be used in similar cases as my combination: low-falling NT plus low-rising NT." His first example - *He's not an 'ordinary 'monkey* - is a most improbable utterance as F + R.

*Allen*¹⁷

On p. 86 of 'Living English Speech' we find: "In some form or another the emphatic pattern of Tune II is extremely common in English, as we so often make remarks which carry with them some implication which is not put into words but is quite clearly expressed by the final rising intonation." Of the illustrative examples I select *All children have to do as they're 'told* (surely F + R?) and the

¹⁵ An Outline of English Phonetics.

¹⁶ Intonation of Conversational English.

¹⁷ Living English Speech.

immediately following *`This material's very 'poor* (ambiguous but probably intended as FR: F + R could stand in the sense 'This material too' or 'This material, for example').

Allen appears to be guilty of the much rarer solecism of confusing FR and F + R with ordinary 'Tune II'. Thus he gives (p. 55) *'Don't 'hurt the 'poor thing* and (p. 97) *'Have a 'heart*. The first (surely *`hurt?*), hard though I have tried to contextualize it, I still find little short of ludicrous; and the second strikes me as an improbably casual intonation for a form of words regularly associated with entreaty, and which is in my experience normally heard as *`Have a 'heart*.

*Palmer*¹⁸

It is a relief to acknowledge that *Palmer*, though he virtually ignores problems of stress and excludes FR from questions, is, within these limitations, rarely in error. It is a little hard, therefore, to find that for his perspicacity he was actually taken to task by *Schubiger* (in an earlier article)¹⁹ for marking differently sentences with identical intonations and using the same notation for dissimilar sentences. The examples she quoted in this connexion are, in my view, all perfectly sound.

*MacCarthy*²⁰

The places where I find myself in disagreement with *MacCarthy* are mostly cases of his indicating F + R where I feel that FR is required. On page 58 of his 'English Conversation Reader' we find *That 'certainly 'is the 'quietest 'time*. Now the point surely is that the time in question is the *quietest*, but in practice unsuitable. The prominence given to 'time' is unnecessary in this context. The sentence is not, it seems to me, of the same order as my example 16: as marked by *MacCarthy* it indicates that when everyone is asleep things are very quiet indeed, which is more or less self-evident, and *fails* to indicate the contrast between that high degree of quiet – apparently so conducive to prolonged reading – and its uselessness in practice. Again, on page 64 we find *`that wouldn't 'do* in a context which clearly indicates a contrast between 'that' (an unsuitable

¹⁸ English Intonation with Systematic Exercises.

¹⁹ *Schubiger*, 1946, p. 132, n. 5.

²⁰ English Conversation Reader.

example that has been suggested) and the correct examples that have still to be found. Similarly, on page 65 we find *'I don't 'know*, which in the context obviously implies that one would need to ask someone else. There are one or two other similar examples.

*Lee*²¹

In making the distinction between TONE and SEQUENCE, *Lee* states that in the case of the former the rise can begin at once after the fall. With this I agree: but he goes on to say that it need not begin at once "unless there is danger of confusion". I find it difficult to accept that the danger of confusion has very much to do with the shape of a spontaneous utterance. I think we must just accept that confusion can, and does, sometimes occur. In the same way, when *Lee* states that in *She doesn't speak to 'anybody she meets along the road* the rise must start within *anybody*, whereas in *She doesn't speak to 'anybody she meets along the 'road* it must not, I fear that, desirable though this compulsion might be, he is rather overstating the case.

According to *Lee*, "sometimes the distinction of usage" is not to be seen "so that *'Do make haste* and *'Do make 'haste* appear to be equivalent, as do *'That doesn't matter* and *'That doesn't 'matter*". It is not entirely clear what *Lee* means by this. From what he has said earlier, we may presume that he considers that there is no important distinction of meaning, and that therefore these sentences do not need to be kept apart in pronunciation. What kind of pronunciation do we, then, in fact, get? It is, however, perfectly clear what all four sentences would mean if used, and the fact that *'Do make haste* is an extremely improbable utterance does not alter this. *'That doesn't matter* is no doubt much more common than *'That doesn't 'matter*, but the latter may mean, for instance, "Even that is no cause for worry". Similarly, when *Lee* says that *Kingdon's* sentence *Not 'all the doctors in the world could 'cure him* might equally well have FR (i.e. *'all*) without change of meaning, I am entirely unable to agree. As marked – that is, with F + R – it means that the combined attentions of the world's doctors would be (or were) unavailing, whereas with FR it might be the boast of one of the few doctors who *could* cure him.

The role of FR in questions, about which *Lee* shows some reserve, has already been dealt with in the appropriate place.

²¹ Fall-Rise Intonations in English.

*Schubiger*²²

It is impossible to get a clear idea from what *Schubiger* says in her recent paper how far she considers that the various patterns she sets up differ or do not differ in their exponents. Though she mostly seems to imply that Types 1–3 are practically indistinguishable, it is only fair to add that she seems vaguely aware of the pitch differences I have spoken of in my own descriptions of FR and F + R. In dealing with the differences between her Type I (my F + R) and her Type 2 (my FR), she says, in a very obscure passage: “*There were* `seven of `them. *I’ve for* `gotten it `now. *We disa* `gree with each `other might be both 1 and 2, though 2, which is an emphatic variant of Tune II, has probably a wider pitch-range. One can also bring out the difference by saying the sentence in the same mood and dropping the last words.” Now by ‘bring out the difference’, coupled with her failure to indicate any preference, *Schubiger* presumably implies that the examples given might be found with either pattern. On the other hand, her use of *though* does not help this interpretation. If it is implied that both FR and F + R are to be found on the substrates given, the examples seem particularly ill chosen. *I’ve forgotten it now* is most unlikely to be FR, and *We disagree with each other* equally unlikely to be F + R. The latter case would give us a meaning something like “What we do with each other is to disagree”.

Schubiger’s Type 3 examples are almost all instances of my F + R, and *Lee* likewise suggests that they might be considered as belonging to a sub-class of his Sequence. The ‘odd man out’ in *Schubiger*’s examples is `That’s `funny. This substrate, which *Allen* gives as ‘That’s `funny, is in my speech usually associated with FR, that is, `That’s funny. I do not understand how *Schubiger* decides when a sentence belongs to Type 3. I presume that the identification of the rise as the nucleus depends on the ‘existence’ of a ‘variant’ with a rise without a preceding fall, but must this be the *only* variant, for instance? *Lee* evidently understands the technique better than I, since he produces an example of his own – *It* `wasn’t in the `middle – which, however, I can see no good reason for allotting to Type 3.

²² Again: Fall-Rise Intonations in English.

Summary

I may summarize my findings as follows. Over a wide range of English sentence-types it is necessary to differentiate between two closely similar falling-rising intonation patterns, one of which may be regarded as unitary, the other (F + R) as binary. I prefer to restrict the term 'fall-rise' to the unitary pattern. As abstracted from large numbers of utterances, the two patterns may be formally distinguished by criteria of pitch and rhythm, but their exponents shade off into each other in such a way that in any given instance it may be difficult or impossible to recognize by ear alone which pattern is to hand. These formal differentiae are found, however, to correlate with 'meanings' which are difficult to summarize, but which for each pattern seem broadly reducible to an unity over a high proportion of all utterances so characterized. These meanings, once appreciated, may be used diagnostically both in listening to ambiguous data and for purposes of textual transcription. Existing textual material shows a great deal of inconsistency in this respect.

I am fully conscious of the deficiencies of this study. More particularly, some of my readers may well find my descriptions of the 'meanings' of FR and F + R of little value to them. In case, however, there are those who would challenge the very propriety of attempting such descriptions, I must emphasize yet again that the formal criteria for the differentiation have absolute priority, and that such hints of a more notional kind as I feel able to give are introduced only in order to make the distinction more credible and to enable others to re-examine the phonic data for themselves. Eighty years ago *Henry Sweet* found it necessary to write: "The whole relation of tone to language has as yet been only imperfectly studied", and his words remain essentially true today. If the study of this relationship is not to remain for ever the Cinderella of the linguistic sciences, then intonation must put its own house in order; and in our efforts to facilitate and, if possible, accelerate that process, we shall need all the resources that we can muster.

Zusammenfassung

Für sehr viele englische Satztypen ist es notwendig, zwischen zwei ziemlich ähnlichen fallenden-steigenden Intonationsformen zu unterscheiden. Die eine ist einheitlich, die andere zweigliedrig. Der Ausdruck *fall-rise* wird nur für die einheitliche Form angewandt. Beide Formen können durch Tonhöhe und Rhythmus unterschieden werden, aber sie gehen in Grenzfällen so ineinander über, daß es gegebenenfalls schwierig

oder unmöglich ist, allein durch das Ohr zu unterscheiden, welche Form vorliegt. Die formalen Unterschiede hängen mit «Bedeutungen» zusammen; diese sind schwer zusammenzufassen. Es scheint aber, daß diese Bedeutungen sich für jede der beiden Formen überhaupt auf eine Einheit zurückführen lassen. Wenn man diese Bedeutungen erst einmal bemerkt hat, kann man sie zur Diagnose verwenden sowohl beim Hören unbestimmter Rede als auch bei Textumschriften. Die bisher vorliegenden Texte stimmen in dieser Hinsicht wenig überein.

Résumé

Pour bien des genres de la phrase anglaise, il faut distinguer entre deux mélodies de type descendant-ascendant assez semblables. L'une est unitaire, l'autre binaire. L'expression *fall-rise* sera employée uniquement pour la mélodie unitaire. Les deux mélodies se distinguent par la hauteur et par le rythme: cependant, dans les cas limites, elles empiètent l'une sur l'autre à tel point qu'il est souvent difficile, ou même impossible à l'oreille seule de décider de quel type il s'agit. Aux différences formelles correspondent des «sens»; ceux-ci sont difficiles à résumer. Il semble cependant, que, pour chacun des deux types, ces sens se laissent dans l'ensemble réduire à l'unité. Si l'on a une fois pris note de ce sens, on peut l'employer comme critère aussi bien pour le discours ambigu à l'oreille que pour les transcriptions écrites. De ce point de vue, il y a peu d'accord dans les textes dont nous disposons jusqu'à présent.

References

- Allen, W. Stannard: *Living English Speech* (Longmans, London 1954).
 Armstrong, L. E. and Ward, I. C.: *A Handbook of English Intonation*. 2nd ed. (Heffer, Cambridge 1931).
 Hockett, C. F.: *A Manual of Phonology* (Waverly Press Inc., Baltimore 1955).
 Jassem, W.: *Intonation of Conversational English* (Nakkadem Wrockawskiego Towarzystwa Naukowego, Wrocław 1952).
 Jones, D.: *An Outline of English Phonetics*. 6th ed. (Heffer, Cambridge 1947).
 Kingdon, R.: *The Teaching of English Intonation*. *English Language Teaching* 2: 118–120 (1948).
 Id.: *The Semantic Functions of Stress and Tone*. *English Language Teaching* 3: 181 (1949).
 Lee, W. R.: *Fall-Rise Intonations in English*. *English Studies* 37: 62–72, 160–162 (1956).
 MacCarthy, P. A. D.: *English Conversation Reader* (Longmans, London 1956).
 Palmer, H. E.: *English Intonation with Systematic Exercises* (Heffer, Cambridge 1922).
 Schubiger, M.: *Again: Fall-Rise Intonations in English*. *English Studies* 37: 157–160 (1956).
 Id.: *The Role of Intonation in Spoken English* (Fehrsche Buchhandlung, St. Gallen 1935).
 Id.: *Intonation – Word-Order – Provisional It*. *English Studies* 27: 129–141 (1946).
 Scott, N. C.: *An Experiment on Stress Perception*. *Le Maître Phonétique* 1939: 44.
 Sweet, H.: *A Primer of Spoken English*. 4th ed. (Clarendon Press, Oxford 1932).
 Trager, G. L. and Smith H. L., Jr.: *An Outline of English Structure* (Norman, Oklahoma 1951).

Author's address: Alan E. Sharp, School of Oriental and African Studies, London W.C.1 (England)

Aus dem Phonetischen Laboratorium des Sprachwissenschaftlichen Instituts der
Akademie der Wissenschaften in Ungarn

Neuere Untersuchungen über die ungarischen Affrikaten

Von LUDWIG HEGEDŰS, Budapest

1. In den dreißiger Jahren wurde in Ungarn anschließend an *Anton Horgers* Artikel «Was ist eine Affrikate?» eine Affrikaten-debatte in den Spalten der Zeitschrift *Magyar Nyelv* eingeleitet¹. Nachdem *Horger* die phonetische Natur der ungarischen Affrikaten untersucht hatte, gab er seine frühere Meinung auf, daß nämlich die Affrikate die Verbindung eines Verschußlautes und eines homorganen Engelautes sei, und kam auf Grund seiner Beobachtungen, noch mehr auf Grund lautgeschichtlicher Angaben zu dem Resultat, daß die Affrikaten «Einzellaute sind und die Dauer ihrer Bildung auch nicht länger ist als die der kurzen Einzelkonsonanten». Im weiteren: «Wir bilden und hören nur einen Laut... in dem die akustischen Wirkungen des Verschußlautes und des... Engelautes miteinander vermengt sind.» In meinen Artikeln «Zur Frage der Affrikaten²» und «Phonetische Randbemerkungen³» kam ich auf Grund von experimentalphonetischen Aufnahmen und Dauermessungen im Gegensatz zu *Horgers* Feststellungen zum Resultat, daß die Affrikaten *in artikulatorischer Hinsicht* keine einheitlichen Laute sind: «a) bei der Bildung der Affrikaten können zwei Momente beobachtet werden, und das eine weist die charakteristische

¹ *Horger*, 1935, S. 210–218; 1936a, S. 100–103; 1936b, S. 259–260; 1942, S. 285–286. *Hegedűs*, 1936a, S. 17–20; 1936b, S. 189–191. *Madzsar*, S. 103–106. *Kelemen*, S. 119. *Hartnagel*, S. 319–321. *Lazicius*, 1937, S. 76–82. *Bakó*, 1942a, S. 167–172; 1942b, S. 357–358; 1937, S. 35–57.

² *Hegedűs*, 1936a, S. 17–20.

³ *Id.*: 1936b, S. 189–191.

Bildungsweise der Verschlußlaute, das andere die der Engelaute auf; b) die Zeitdauer der Bildung der Affrikaten ist beträchtlich länger als die der kurzen Konsonanten¹; der Verschluß und die Enge werden nicht gleichzeitig gebildet, sondern folgen zeitlich aufeinander, und da ihre Bildung an derselben oder an einander sehr nahe liegenden Stellen erfolgt, passen sich ihre Artikulationsmomente einander an². Auch hielt ich für wesentlich, das Artikulationsmoment und das *akustische Moment* auseinanderzuhalten. Akustisch hat nämlich die Klangfarbe der Affrikaten einen ziemlich breiten Spielraum. Der Eindruck, den der Laut *c* (*ts*) macht, wird im allgemeinen einheitlicher aufgefaßt als der des *dz*: «Die Affrikaten des ungarischen Wortes *cica* ‚Kätzchen‘ können von einzelnen eher als Einzellaute empfunden werden (da *c* mit einem Buchstaben bezeichnet wird) als z. B. die der Wörter *bodza* ‚Holunder‘, *findzsa* ‚Tasse‘, *lándzsa* ‚Lanze‘ usw. Hier kann jeder, auch ohne scharfe Ohren zu haben, den Doppelcharakter der Affrikaten erkennen³.»

Im Verlaufe der Debatte wurden manche Fragen berührt, doch selbst nach der Zusammenfassung der Debatte durch *Laziczzius* wurden die Begriffe nicht befriedigend geklärt⁴. Der Grund hierfür war, daß *Laziczzius* keine einschlägigen phonetischen Untersuchungen angestellt und die Feststellungen der Debattierenden nicht mit Messungen und phonetischen Untersuchungen kontrolliert hatte und seine Meinung ausschließlich auf eine Arbeit *Belgeris*⁵ gründete, indem er dessen Schlußfolgerung annahm: «Die Affrikate ist ein gemischter Laut (,un phonème mixte‘), der ganz eigenartig gebildet wird. Die Bildungsweise der Affrikate steht in der Mitte zwischen der Bildungsweise des Verschlußlautes und des Engelautes. Sie ist mit keinem dieser Laute identisch, auch keine Summe der beiden. Sie ist nicht einmal ein einfacher Übergang von einem zum anderen, sondern die Verschmelzung der beiden, die oft vollkommen, noch öfter aber unvollkommen ist⁶.»

¹ *Hegedüs*, 1936b, S. 191.

² *Id.*: 1936a, S. 19.

³ *Id.*: 1936b, S. 190.

⁴ *Laziczzius*, 1937, S. 76–82; 1944, S. 78–84.

⁵ *Belgeri*, 1929.

⁶ *Laziczzius*, 1937, S. 81. Auf die Unhaltbarkeit von *Belgeris* «phonème mixte» hat auch *J. Forchhammer* (*Arch. Néerl. Phon. Exp.* 17; 12, 1941) hingewiesen: «Wenn z. B. *Belgeri*... von einer Vermischung von Verschlußlaut und Engelaute spricht, so ist dazu zu bemerken, daß eine solche Vermischung überhaupt nicht stattfinden kann; denn, wenn es auch möglich ist, während eines Verschlusses eine Enge zu bilden, so kommt diese Enge jedoch erst *nach* der Lösung des Verschlusses phonetisch zur Geltung.»

2. Durch die Auffrischung der *Belgerischen* Theorie wurde aber die Frage bei weitem nicht geklärt, die Verwirrung ist vielmehr noch größer geworden. Dies erhellt auch aus *E. Bakós* Affrikatenstudie¹: «In phonetischer Hinsicht können wir die Sache selbst nach dem reichen Material und den Feststellungen *Belgeris* nicht für abgeschlossen halten, denn die Frage ist noch offen geblieben, was die Affrikate von phonetischem Gesichtspunkt aus ist². *Bakó* sucht die rätselhafte Natur der Affrikaten auf Grund von Kymogrammen und Palatogrammen aufzuhellen, seine Aufnahmen legt er aber willkürlich aus. Seine Kymogramme zeigen deutlich das Moment *Verschuß + Enge*, von seinen Palatogrammen weisen *dz*, *gy* (*d'j'*) und *ty* (*t'z'*) vollkommene Verschußbildung, *c* (*ts*), *cs* (*tš*) und *dz* verminderte Engebildung auf, und trotzdem lautet seine Schlußfolgerung für die Artikulation folgendermaßen: «Bei der Bildung der Affrikaten lassen sich drei Momente beobachten: ein engeverminderndes, ein engeerhaltendes und ein engeerweiterndes Moment³.» Ihrem Verschußelement schenkt er kaum eine Beachtung, und ihren Platz weist er einfach in der Gruppe der Engelaute an: «Die ungarischen Affrikaten sind mit stark verminderter Zungenenge gebildete *Engelaute*.» Es liegt auf der Hand, daß die Zuteilung der Affrikaten zu den einfachen Engelaute nicht annehmbar ist, da dies jeder Art von subjektiver Beobachtung und experimentalphonetischer Analyse vollkommen widerspricht. Deshalb weist *Horger* eine derartige Auslegung *Bakós* zurück⁴: «Daß die Affrikaten keine Verschußlaute, sondern Engelaute sind (die eine schmalere Enge als die anderen haben), kann ich keineswegs zugeben.» Und: «Nie, kein einziges Mal habe ich beobachtet, daß ich nicht einen Verschuß, sondern eine mehr oder minder schmale Enge gebildet hätte.»

Von dieser Kritik beeindruckt zog dann *Bakó* seine Engelaute-theorie zurück und teilte die Affrikaten in die Gruppe der Halbverschußlaute («semiexplosivae») ein – nicht anders als *Belgeri* –: «Die Kategorie der Halbverschußlaute steht in der Mitte zwischen der Gruppe der Verschußlaute und der der Engelaute, so hat sie auch im System der ungarischen Verschußlaute zwischen den

¹ *Phonetische Studien* (ungarisch). Debrecen 1937.

² a. a. O., S. 38.

³ a. a. O., S. 57.

⁴ *Horger*, 1942, S. 286.

beiden Gruppen ihren Platz¹. Die Zuordnung zu den Engelaute hat also *Bakó* am Ende selbst fallenlassen, trotzdem taucht die «Engelauttheorie» sowohl in der ungarischen als auch in der ausländischen Fachliteratur noch immer auf.

Die Beurteilung der Affrikaten ist noch immer sehr unsicher, und diese Unsicherheit kann im Universitäts- und Hochschulunterricht wie auch in den Debatten über die ungarische Rechtschreibung verfolgt werden. Auf die Widersprüche der Notizenhefte und Handbücher habe ich bereits an einem anderen Ort hingewiesen. Aus dem Gesagten geht klar hervor, daß die Frage der Affrikaten noch nicht in zufriedenstellender Weise geklärt ist und die Sprachforscher durch dunkle Erklärungen nicht befriedigt werden können. Die Unsicherheit der Beurteilung hat nicht nur bei uns, sondern auch im Kreise ausländischer Phonetiker und Sprachforscher lebhaft Auseinandersetzungen entfesselt, und es wurden an vielen Orten nützliche experimentalphonetische Analysen angestellt, um die eigenartige Natur der Affrikaten zu erkennen. Nach *Belgeris* Analysen habe ich die Ergebnisse meiner experimentalphonetischen Untersuchungen bezüglich der ungarischen Affrikaten in meinem Aufsatz «Die Natur der ungarischen Affrikaten²» veröffentlicht. *E. Richter* hat in ihrer Studie «Die italienischen č- und š-Laute³» italienische Affrikaten von vielen Seiten her analysiert. *J. Forchhammer* hat *E. Richter* kritisiert⁴ und in seinen Arbeiten «Kern- und Wendepunkt der Sprechwissenschaft⁵», «Zum Affrikaten-Problem⁶» wie auch in seiner «Laletik⁷» manche Einzelheiten der Affrikatenfrage geklärt. An der Debatte beteiligte sich auch *D. Gerhardt*⁸, und auch in Amerika kam es zu Debatten über die Affrikaten⁹. Unlängst hat sich der hervorragende tschechische Phonetiker *B. Hála* mit dieser Frage eingehend beschäftigt¹⁰.

Nach den ausgetragenen Debatten, hauptsächlich aber auf Grund experimentalphonetischer Untersuchungen, beurteilen wir

¹ *Belgeri*, S. 170.

² *Hegedüs*, 1939, S. 97–102.

³ *Richter*, S. 1–38.

⁴ *Forchhammer*, 1941, S. 9–20.

⁵ *Id.*, 1951a, S. 23–36.

⁶ *Id.*, 1953, S. 403–408.

⁷ *Id.*, 1951b.

⁸ *Gerhardt*, 1948, S. 95–96; 1952, S. 57–76.

⁹ *Martinet*, S. 116–122.

¹⁰ *Hála*, S. 77–93.

vielleicht die Affrikaten richtiger; wir können auch sehen, was die Fäden dieser Frage verwirrt hat, warum wir so langsam vorwärtsgekommen waren.

3. Bei der richtigen Beurteilung der Affrikaten hatte ohne Zweifel die größte Verwirrung der Umstand hervorgerufen, daß die Debattierenden und die Forscher ihre von verschiedenen Gesichtspunkten aus gemachten Beobachtungen nicht auseinandergehalten hatten. Mit Recht hat *J. Forchhammer* im Zusammenhang mit *E. Richters* Untersuchungen auf diesen Grundfehler hingewiesen, die (d. h. *Richter*) – wie auch viele andere – den zusammengesetzten Charakter der Affrikaten sowohl in artikulatorischer als auch in akustischer Hinsicht nachgewiesen und trotzdem die Schlußfolgerung gezogen hat, daß sie «einheitliche Laute» sind: «Wenn die von *E. Richter* angeführten Forscher, trotzdem sie selber den zusammengesetzten Charakter der Affrikaten nachgewiesen haben, doch fast ausnahmslos zu dem Ergebnis kommen, die Affrikaten seien ‚einheitliche Laute‘, so dürfte das in erster Linie auf einer Verwechslung der beiden Begriffe ‚Sprachlaut‘ und ‚Phonem‘ beruhen¹.» Auch die bei uns ausgetragene Affrikatenpolemik hatte ohne Zweifel den Hauptfehler, daß artikulatorische und phonematische Gesichtspunkte bunt miteinander vermengt waren und daß die Debattierenden je nach ihrer Einstellung einmal dem artikulatorischen, ein anderes Mal aber dem phonematischen Gesichtspunkt den Vorzug gegeben hatten.

Heute sehen wir bereits klar, daß wir die Affrikaten nach drei Gesichtspunkten zu analysieren haben. Wir müssen untersuchen 1. von der *physiologischen Seite* her den genauen Ablauf der Artikulation, 2. *von akustischem Gesichtspunkt aus* das Klangbild und den auf die Zuhörer gemachten akustischen Eindruck, endlich 3. das diesbezügliche Urteil der Glieder der Sprachgemeinschaft, den *Phonemwert* der Affrikaten.

4. Ich untersuchte die einzelnen Momente der Artikulation der Affrikaten an Palatogrammen, Kymogrammen, Oszillogrammen, an Bildern, die mit automatischem Tonhöhenschreiber gemacht wurden, und endlich an Schalldruckaufnahmen.

a) Die *Palatogramme* geben Aufschluß über den Umfang und die Form der Berührungsfläche von Zunge und Gaumen, sie geben aber naturgemäß kein Bild vom zeitlichen Verlauf der Artikula-

¹ *Richter*, S. 10–11.

tionsmomente. Meine Palatogramme zeigen, daß in meiner Aussprache bei der Bildung ungarischer Affrikaten immer Verschlußbildung vorhanden war. Der Verschluß der Laute *c* (*ts*) und *dz* wird in alveolarem, der der Laute *cs* (*tʃ*) und *dzs* (*dʒ*) in postalveolarem und der der Laute *ty* (*t'ɟ*) und *gy* (*d'j'*) in postalveolarem bzw. palatalem Gebiet gebildet. Die Verschlußbildung erfolgt immer etwas mehr nach hinten als bei *t* (*t*), weil sie sich an die Stelle des folgenden *s*-Elementes anpaßt. Der Verschluß ist also ein sehr wichtiges Moment der Affrikatenbildung, und unseres Erachtens kann diese Tatsache bei ihrer Beschreibung nicht übersehen werden. Die Berührungsfläche von Zunge und Gaumen ist je nach den Affrikatentypen und der Zeitdauer verschieden. Verhältnismäßig am kleinsten ist sie bei *dz* und *c* (*ts*)¹, etwas größer bei *dzs* (*dʒ*) und *cs* (*tʃ*) und naturgemäß am größten bei der langen palatalisierten Gruppe, bei den Lauten *gy* (*d'j'*) und *ty* (*t'ɟ*). Bei der Bildung dieser beiden letzteren Laute decken sich Palatalisation und Affrikation. Das spirantische Element entsteht zusammen mit der Sprengung der auf einer großen Fläche erfolgten Verschlußbildung und ist eine notwendige begleitende Geräuscheinung. In den phonetischen Einteilungen werden sie (*gy*, *ty*) gerade deshalb bald der Gruppe der palatalisierten Laute, bald den Affrikaten zugeteilt. An Palatogrammen kann auch noch beobachtet werden, daß die Berührungsfläche der stimmhaften Affrikaten kleiner ist als die der entsprechenden stimmlosen, weil die Verschlußbildung mit kleinerer Muskelspannung erfolgt.

Meine Palatogramme zeigen ebenso wie die Gaumenbilder von *Ž. Gombocz*², *J. Balassa*³, *Rousselot*⁴, *B. Csüry*⁵ und *B. Vass*⁶ bei der Artikulation eine Verschlußbildung. Sie weichen aber stark von den Palatogrammbildern von *E. Bakó*⁷ ab, auf dessen Palatogrammen *c* (*ts*) eine Enge, sein stimmhaftes Korrelat *dz* aber – das mit viel schwächerer Muskeltätigkeit gebildet wird – einen vollen Verschluß zeigt! Wir halten es für natürlich, daß auch die Affrikaten bei der Aussprache der Glieder einer Sprachgemeinschaft einen

¹ Hegedüs, 1941/43, S. 64–73; 1955, S. 181–214.

² Gombocz, S. 193–204.

³ Balassa, S. 85.

⁴ Rousselot, S. 609.

⁵ Csüry, S. 67.

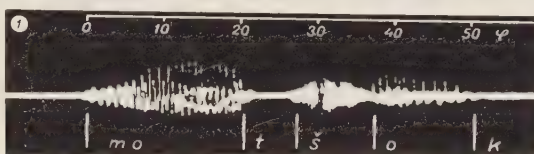
⁶ Vass, S. 40–43.

⁷ Bakó, 1937, S. 48–56.

gewissen Schwankungsbereich haben¹, die Bildung ohne Verschuß bezeichnet aber offenbar nur die äußerste Grenze dieses Schwankungsbereiches, und nicht die verbreitetste Bildungsweise. So etwas kann hauptsächlich bei sehr schnellem Sprechen vorkommen, wo die auf die Artikulation verwendete Zeit in hohem Maße verkürzt wird und so die Reduktion der Zeitdauer der einzelnen Elemente hochgradig sein kann. Der Umfang der bei der Verschußbildung entstehenden Berührungsfläche wächst in geradem Verhältnis zum Anwachsen der Zeitdauer², eben deshalb können wir bei der langen Variante der Affrikaten sowohl das Verschußelement als auch das Engeelement deutlich wahrnehmen und unterscheiden.

b) Von *Kymogrammen* können wir auch das Nacheinander der Artikulationsmomente der Affrikaten ablesen³: *zuerst tritt das Bild des Verschlusses hervor, danach das des Engeelementes*. Die beiden kann man voneinander abgrenzen, ihre Dauerverhältnisse sind aber je nach Affrikatentypen verschieden.

c) Auf *Oszillogrammen* können wir die Momente der Bildung noch eingehender studieren. Die Kathodenröhre registriert nämlich die Sprechlaute frei von Trägheit.



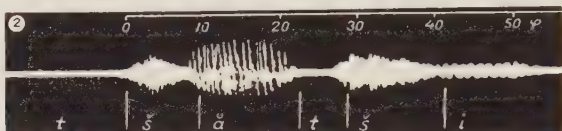
cs (tš): Die erste Abbildung zeigt das Oszillogramm des ungarischen Wortes *mocsok* «Schmutz». Nach den Wellen des Lautes *o* sehen wir das Bild des Verschußelementes, aus dem sich das Geräuschelement *š* von hoher Frequenz mit allmählich steigender, dann abnehmender Amplitude hervorhebt. Der reduzierte Verschußlaut (Dauer: 5 φ) und der Engelaut (10 φ) werden also nicht gleichzeitig und miteinander vermengt gebildet – wie es *Belgeri*, *Horger*, *Laziczius* und noch viele andere gemeint haben –, sondern in zeitlichem Nacheinander. Es trifft also nicht zu, was *Laziczius* sagt: «das ganze Registrat (zeigt) eine unanalysierbare Einheit, die

¹ Auch in der Arbeit «Polskie Affrykaty» von *Maria Dluska* können wir eine gewisse Schwankung bei der Artikulation beobachten. Von den 9 Palatogrammen der polnischen Laute *ts* und *tʃ* weisen sieben volle, zwei aber unvollkommene Verschußbildung auf.

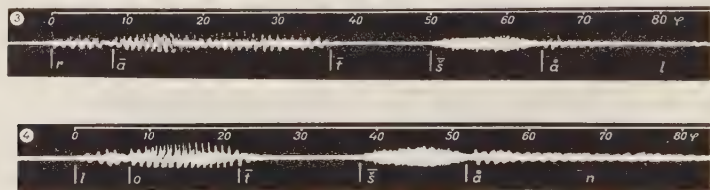
² *Hegedüs*, 1941/43, S. 64–74; *Stetson, Hudgins and Moses*, S. 52–61.

³ *Meyer und Gombocz*, 1909; *Hegedüs*, 1936a, S. 17–20; 1939, S. 97–102.

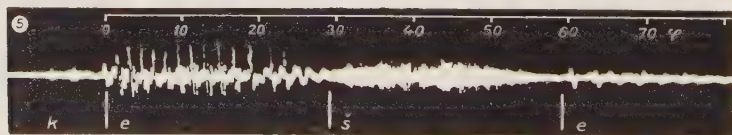
nicht in ihre Elemente zergliedert werden kann. Es kann also keineswegs gesagt werden, wie lang der Kurventeil des Verschlußelementes anzusetzen ist und wo der Kurventeil des Engeelementes beginnt¹.» Im Gegenteil: man kann die zwei Elemente genau voneinander abgrenzen.



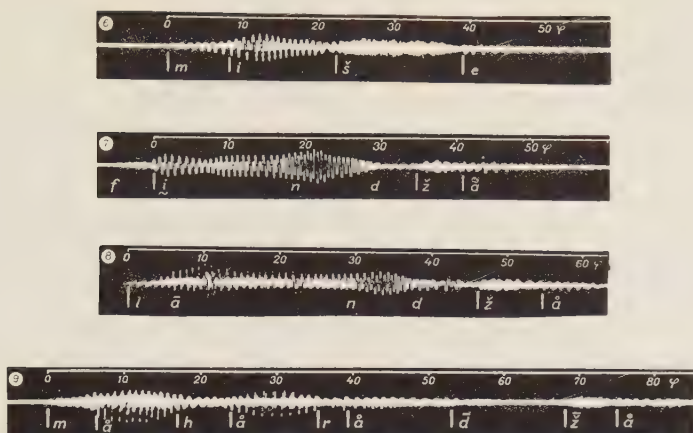
Auf Abbildung 2 (*csacsi* «Eselchen», «Eselfohlen») sehen wir einen ähnlichen Fall. Hier hebt sich auch das zischende Lautelement des *cs* (*tʃ*) am Anfang des Wortes plastisch hervor. Auf Oszillogrammen, die von langen Affrikaten mit großer Geschwindigkeit gemacht wurden (Abb. 3: *ráccsal* [*rātʃál*] «mit Gitter», Abb. 4: *locsan* [*lotʃán*] «platscht»), treten die zeitlich aufeinander-



folgenden Affrikatenelemente noch anschaulicher hervor. Daß das zweite Element der Affrikate in der Tat dem Laut *ʃ* gleich ist, wird auch durch akustische Teiltonanalysen bestätigt. Meine Versuche mit Oktavsieben haben gezeigt, daß sowohl das *ʃ*-Element der Affrikaten als auch die *ʃ*-Laute die charakteristische Frequenzstelle beim Stand I. 2400–4800 Hz bzw. II. 3200–6400 Hz des Gerätes hatten. Außerdem kann uns die Vergleichung mit den Oszillogrammen der Wörter, die einen *ʃ*-Laut enthalten, von ihrer Identität überzeugen (Abb. 5: *kese* [*keʃe*] «falb»; Abb. 6: *mise* [*miʃe*] «Messe»).

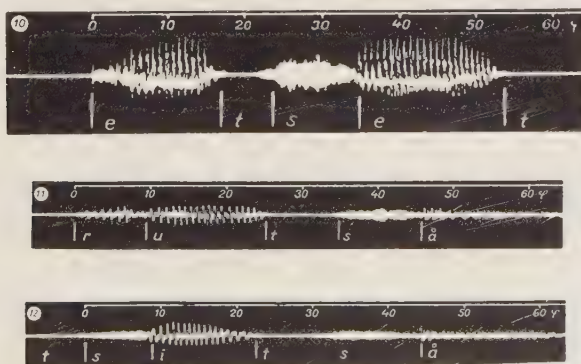


¹ Laziczius, 1944, S. 80–81.

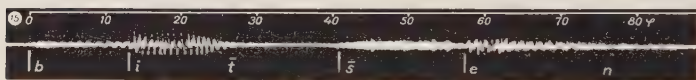
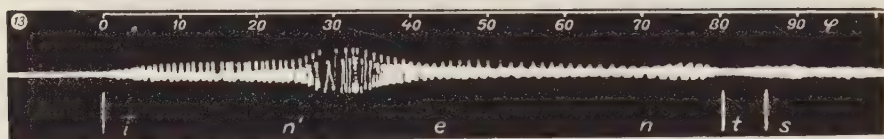


dzs (dʒ): Die artikulatorische und akustische Zweigliedrigkeit des stimmhaften Korrelats des *cs* ist in den Oszillogrammen ebenfalls gut wiedergegeben. Die Amplitude der Wellen des mit reduzierter Muskeltätigkeit gebildeten Verschlusselementes (*d*) ist klein, und seine Dauer ist kurz (Abb. 7: *findzsa* [findžä] «Tasse» und Abb. 8: *lándzsa* [ländžä] «Lanze») bzw. lang (Abb. 9: *maharadzsa* [mähârädžä] «Maharadscha»); es ist auch vom Element *ʒ*, das mit Geräuschen von hoher Frequenz gebildet wird, deutlich abgrenzbar.

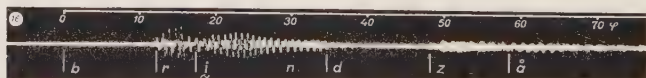
c (ts): Was den zeitlichen Ablauf der Bildung der Affrikate *c* betrifft, so ist er dem der vorigen gleich. Das Zischelement folgt zeitlich immer auf das Verschlusselement, seine Dauer ist länger als jenes und «vermengt» sich nie mit ihm vom Beginn der Bildung an (Abb. 10: *ecet* [etset] «Essig», Abb. 11: *ruca* [rutsä] «Ente», Abb. 12: *cica* [tsitsä] «Kätzchen»). Die Dauerreduktion des Ver-



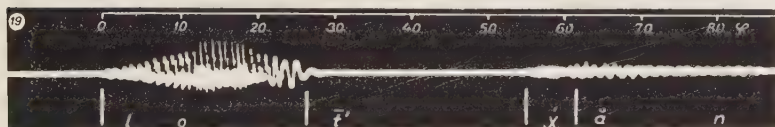
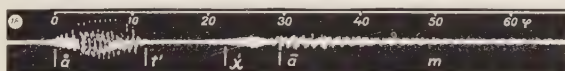
schlußelementes ist im allgemeinen bei den Affrikaten nach *n* am erheblichsten (Abb. 13: *inyenc* [in'ents] «Feinschmecker»). Bei langen Affrikaten wird nicht nur das Verschlusselement, sondern sehr oft auch das Engeelement gelängt (Abb. 14: *ruccan* [rutsân] «einen Gang oder Fahrt tun»: $\bar{t} = 10,1 \varphi$, $\bar{s} = 20,2 \varphi$; Abb. 15: *biccen* [bi̯tsen] «eine kleine wackelnde Bewegung tun»: $\bar{t} = 13,7 \varphi$, $\bar{s} = 16,5 \varphi$). Es ist also klar, daß der Affrikatencharakter nicht ausschließlich durch die Dauerreduktion der Elemente herausgebildet wird, sondern durch andere Reduktionserscheinungen, die infolge der Assimilation (richtiger: Anpassung) entstehen – wovon weiter unten noch die Rede sein wird.



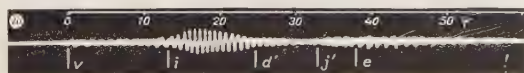
dz: Hier ist die Abgrenzung der beiden Elemente der Affrikate natürlich nicht so plastisch wie bei ihrem stimmlosen Korrelat, aber doch ausgesprochen, denn die Geräusche von hoher Frequenz (I. 4800–9600 Hz bzw. II. 6400–12800 Hz) bezeichnen deutlich die Bildungsphase des zweiten Elementes (Abb. 16: *brindza* [brindzâ] «eine Art Schafkäse»: $\bar{d} = 5 \varphi$, $\bar{z} = 10 \varphi$; Abb. 17: *edzõ* [edzõ] «Trainer»: $\bar{d} = 16 \varphi$, $\bar{z} = 12 \varphi$).



ty (*t'χ'*): Während in der vorigen Gruppe der Affrikaten – zum mindesten in den kurzen Varianten – eher das Verschlusselement hinsichtlich der Dauer reduziert war und das Engeelement dominierte, verhält es sich bei den palatalisierten Lauten eben umgekehrt. Das hauptsächliche Moment der Artikulation ist hier die auf großer Fläche erfolgende *Verschlufbildung*, mit deren Auflösung zwangsläufig Reibegeräusche entstehen. Die Verbindung: Verschluslaut + homorganer Engelaut tritt hier also wie in den oberen Typen auf. Auf Abbildung 18 sehen wir das Oszillogramm des Wortes *atyám* (*ât'χ'ām*) «mein Vater». Das Verschlusselement der Affrikate *ty* (*t'χ'*) kann auch hier vom geräuschhaften Engeelement abgesondert werden. Die Dauer des Verschlusses ist bei der kurzen Variante etwas länger als die des spirantischen Elementes (*t'* = 10,5 φ, *χ'* = 7,5 φ). Bei ihrer langen Variante (Abb. 19: *lottyán* [*lotχ'ân*] «platscht»: *t'* = 28,5 φ, *χ'* = 6 φ) ist die Dauer des Verschlusses 4,7mal größer als die des spirantischen Elementes. Das zweite Element ist, wie gesagt, eine zwangsläufige begleitende Geräuscherscheinung, deren Dauer vom Sprechenden nicht so willkürlich verlängert werden kann wie die des Verschlusses.



gy (*d'j'*): Auch im Korrelat des vorigen Lautes ist der Stimmton im zweiten Element mit Reibegeräuschen durchsetzt. Die charakteristischen Frequenzstellen sind etwas niedriger als z. B. bei *dz*: I. 1200–2400 Hz bzw. II. 1600–3200 Hz (Abb. 20: *vigye!* [*vid'j'e!*] «tragen Sie es!»: *d'* = 8 φ, *j'* = 5 φ).



Die beiden Momente der Artikulation lassen sich auch hier deutlich voneinander abgrenzen.

Die oszillographischen Analysen zeigen also, daß die Affrikaten aus zwei zeitlich nacheinander folgenden Elementen bestehen und daß ihr erstes Element ein Verschlußlaut, das zweite ein homorganer Reibelaut, Engellaut ist. Anhand seiner Oszillogramme meint auch *T. Tarnóczy*¹, daß das zweite Element der ungarischen Affrikaten dem entsprechenden Engellaut gleichgesetzt werden kann, da ihre charakteristischen Frequenzstellen übereinstimmen. Seinen Untersuchungen zufolge war z. B. die charakteristische Frequenzstelle des isoliert gesprochenen Lautes *sz* (*s*) um 6800 Hz herum, das spirantische Element der Affrikate im Worte *cilu* (*tsilu*) «Mieze» zeigte aber einen Wert um 7100 Hz herum. Die charakteristische Formantstelle des isoliert gesprochenen Lautes *s* (*f*) war 5260 Hz bzw. 5800 Hz, das spirantische Element des isolierten *cs* (*tf*) betrug 5200 Hz, 5350 Hz. und 6000 Hz.

Meine Analysen haben also die Ergebnisse meiner früheren mit dem Kymographion angestellten Forschungen bestätigt und stehen in vollem Einklang mit den Untersuchungen von *B. Hála*², der anhand von Oszillogrammen, die er von tschechischen Affrikaten gemacht hatte, feststellte, daß bei den Affrikaten zwei Bildungsmomente: ein Verschlußelement und ein Engeelement zeitlich aufeinander folgen: «On peut aisément distinguer une partie occlusive et une autre fricative; les affriquées seraient donc formées... par deux éléments articulatoires successifs, l'un occlusif, l'autre spirant ou plus précisément ,constrictif‘³.» Auch das hat *Hála* ganz richtig erörtert, daß der Verschluß bei Affrikaten etwas schwächer ist als bei Verschlußlauten und auch das Engeelement etwas unbestimmter ist als bei den echten Zisch- oder Reibelauten. Es ist also offenbar, daß diese Momente hinsichtlich ihrer Bildung typische Reduktionserscheinungen sind, wie es auch *J. Forchhammer* nachdrücklich betont hat: «Die Affrikaten können somit als reduzierte Lalembindungen bezeichnet werden⁴.»

Die bei der Bildung der Affrikaten zutage tretenden zwei Momente können übrigens auch an Oszillogrammen von spanischer und italienischer Rede ähnlich beobachtet werden⁵, wie darauf auch bereits *B. Hála* hingewiesen hat.

¹ *Tarnóczy*, S. 328–329.

² *Hála*, S. 77–93.

³ *Hála*, S. 83.

⁴ *Forchhammer*, 1953, S. 405.

⁵ *Mme De Chaves*, *Problemas de fonética experimental*. 1948, Fig. 56; *Gemelli*, Tafeln XI, XV, XVI und XVII.

d) Mit *automatischem Tonhöhenschreiber* gemachte Aufnahmen gewähren uns auch nützliche Aufschlüsse über die Artikulationsmomente der Affrikaten. Auf diesen Aufnahmen sind Verschlußelement und Engeelement deutlich abgegrenzt, denn während der Verschlußbildung ist die Stille vollkommen (bei stimmlosen Lauten), während der Engebildung entstehen aber Reibegeräusche hoher Frequenz, die bei geeigneter Technik auch auf dem Tonhöhenbild Spuren hinterlassen. Auf Abbildung 21 sehen wir den

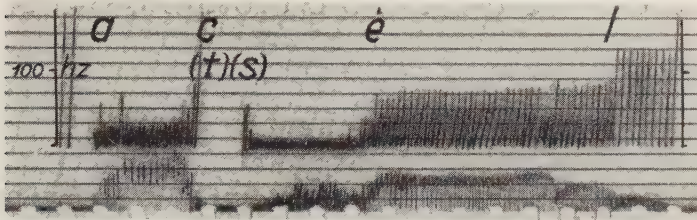


Abb. 21

Tonhöhenverlauf des Wortes *acél* (ätsel) «Stahl» in der Aussageform (oben)¹, und sein gleichzeitiges Oszillogramm (unten). Auf dem Tonhöhenbild beginnt die Darstellung mit dem Vokal *ä* der ersten Silbe, und es weist gegen das Ende hin in den beiden letzteren Perioden längere Lichtstäbchen auf. Hier gehen also die Stimmbänder aus der geschlossenen Stimmtstellung in offene über. Während fünfhundertstel Sekunden zeigt das Tonhöhenbild nichts. Jetzt wird also das Verschlußelement der Affrikate *c* (*ts*) gebildet. Danach erscheint eine Reihe von sehr kurzen Lichtstäbchen auf dem Bild. Dies ist das Zeichen dafür, daß das zweite, d. h. das Engeelement der Affrikate, ein Zischlaut von hoher Frequenz, entstanden ist (*s* = 15 φ). Wir können aber die zwei Phasen der Artikulation nicht nur am Tonhöhenbild, sondern auch am gleichzeitigen Oszillogramm in deutlicher Form beobachten. Unten fällt die Amplitude des Vokals *ä* bei der Bildung des Verschlusses der Affrikate auf die 0-Linie herab, dann wird allmählich die Amplitude der Geräusche breiter und an der Grenze des Vokals *é* (*ē*) wieder schwächer. Es kann selbst mit freiem Auge beobachtet werden, daß das Engemoment ungefähr das Dreifache des Verschlußelementes ist. Es ist also verständlich, daß unser Gehörorgan jedes einzelne Artiku-

¹ Die kurzen Linien bezeichnen hohe musikalische Töne, die längeren tiefere.

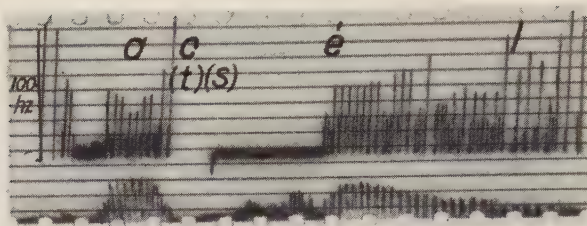


Abb. 22

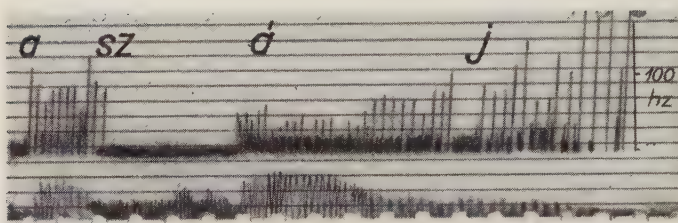


Abb. 23

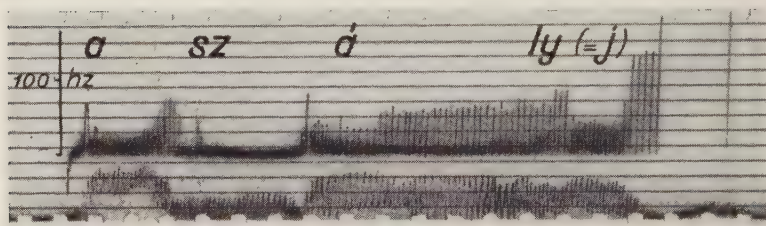
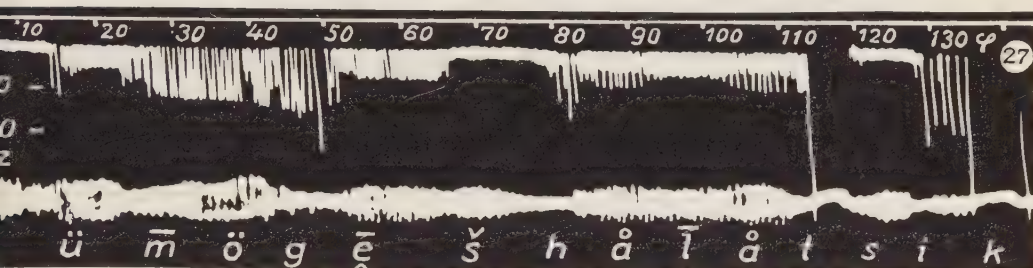
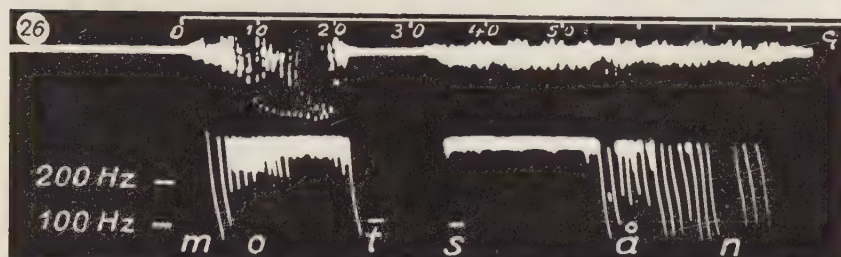
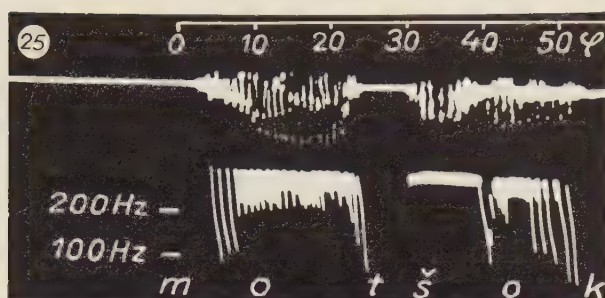


Abb. 24

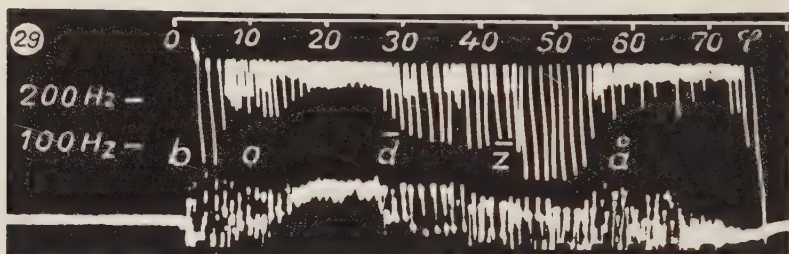
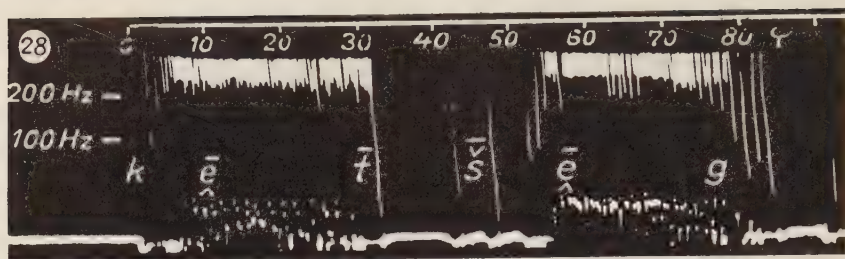
lationsmoment wegen der hochgradigen Dauerreduktion des ersten Elementes nur dann auffassen kann, wenn die Dauer der Affrikate verlängert wird (z. B. *ruccan*) und das sonst reduzierte Verschlußelement in hohem Maße verlängert wird. Ganz ähnlich verhält es sich auch auf Abbildung 22, die das Wort *a cél* (â tsël) «das Ziel» darstellt. Wenn wir diese Wörter mit solchen vergleichen, die den Zischlaut *sz* (*s*) enthalten, sehen wir, daß die Bilder der Engelaute einander ähnlich sind und daß nur an der Dauer einige Abweichung zu bemerken ist (Abb. 23: *a száj* [â sāj] «der Mund», Abb. 24: *aszály* [âsāj] «Dürre»).

Ähnlich verhält es sich auch auf Abbildung 25: *mocsok* (motsok) «Schmutz», 26: *moccan* (motsān) «sich rühren» und 27: *zűmmögés*



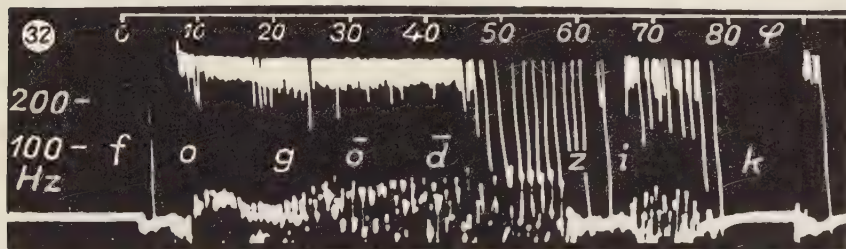
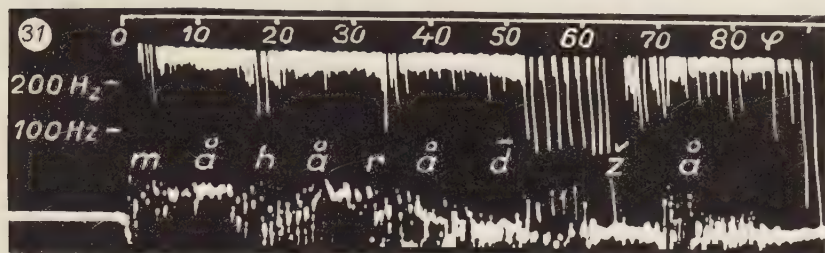
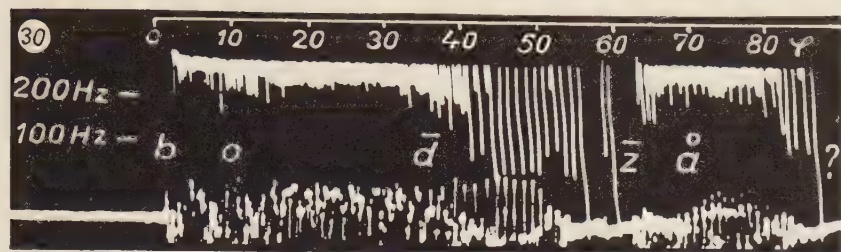
hallatszik (zǖmögēs hālātsik) «ein Summen ist zu hören». Das letztere Teilstück ist aus einer größeren zusammenhängenden Redeeinheit herausgeschnitten, und so zeigt die lange Affrikate des Wortes *hallatszik* – infolge des gesteigerten Sprechtempos – an absolutem Dauerwert keine solche Größe wie in isoliert gesprochenen Wörtern.

Die zwei Phasen der Artikulation der Affrikaten können nicht nur an Aufnahmen mit Mundmikrofon, sondern auch an solchen mit Kehlkopfmikrofon gesehen werden, trotzdem diese Zischgeräusche in solchen Fällen nicht unmittelbar registriert werden können, da diese ja im vorderen Teil der Mundhöhle und nicht in



der Gegend der Kehle entstehen. Abbildung 28 zeigt das Registrat des Wortes *kétség* (*kētšēg*) «Zweifel» mit langer Affrikate. Nach dem Verschlusselement bemerkt man während der Bildung des Zischlautes auch Geräusche von ganz niedriger Frequenz. Abbildung 29 gibt den Tonverlauf und das Oszillogramm des Wortes *bodza* (*bodzâ*) «Holunder» wieder. Beim Verschlusselement der Affrikate fällt auch die musikalische Höhe des Tons allmählich tiefer und erreicht ihre größte Tiefe während der Artikulation des Engelelementes. Das Verhältnis der Dauer des Verschlusselementes und des Engelelementes ist 1,6:1 ($\bar{d} = 16 \varphi$, $\bar{z} = 10 \varphi$).

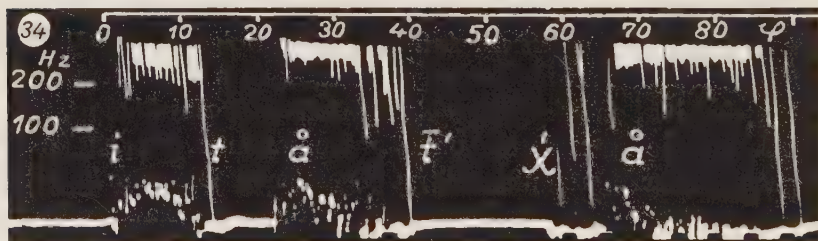
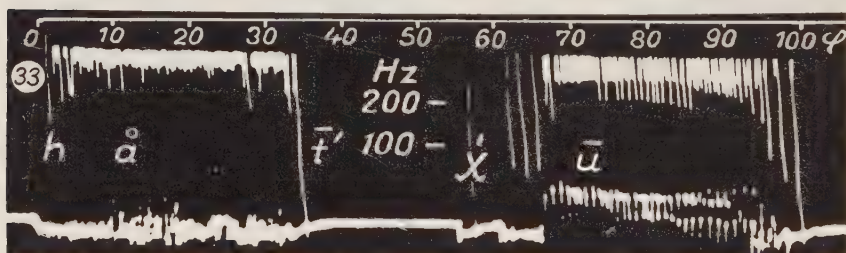
Wie oben bereits erwähnt, kann die Dauer der einzelnen Elemente der Affrikate reduziert sein. Aus den experimentalphonetischen Untersuchungen erhellt aber auch, daß oft auch die Stimmippentätigkeit im zweiten Element der stimmhaften Affrikaten eine gewisse Reduktion aufweist. Abbildung 30 zeigt die Frageform des Wortes *bodza* (*bodzâ*) «Holunder». Am unteren Oszillogramm können wir deutlich sehen, daß die Amplitude der von den Stimmbändern erzeugten Wellen im zweiten Element der Affrikate sehr klein ist und der Laut gegen das Ende hin fast als stimmlos angesprochen werden kann. Daß es wirklich so ist, können wir auch am oberen Tonhöhenbild kontrollieren, wo die Kontinuität der Lichtstäbchen an mehreren Stellen aussetzt. In dieser Hinsicht sind auch



die Abbildung 31: *maharadzsa* (māhârād̄z̄ā) «Maharadscha» und 32: *fogódzik* (fogód̄zik) «sich anhalten» sehr lehrreich; an ihren Oszillogrammen können wir die Reduktion der Stimmlippentätigkeit im zweiten Moment der Affrikate gut verfolgen.

Es erhellt also auch aus Analysen mit dem automatischen Tonhöhenschreiber, daß die *Bildung der Affrikaten nicht einheitlich ist*, also anders als die Artikulation anderer Konsonantengruppen (z. B. Engelaute oder Verschußlaute) im allgemeinen. Die charakteristischen Momente der Verschußlaute sowie der Engelaute sind in ihnen enthalten, aber meistens ist keines von ihnen vollkommen, da sie sich an Bildungsweise einander anpassen: die Stelle des Verschlusses wird ein wenig nach der Enge zu verlagert, der Verschuß hingegen wird nicht ganz gesprengt, nur bis zu der Stellung, die die Engbildung benötigt. Außerdem wird – je nach den verschie-

denen Typen der Affrikaten – Dauer, Stimmlippentätigkeit, Muskelspannung oder irgendein anderes Artikulationsmoment des einen oder des anderen Elementes reduziert. Bei der Artikulation der kurzen Affrikaten vom Typ *c* (*ts*) und *cs* (*tʃ*) ist die Reduktion des Verschlusselementes auffallend, das im allgemeinen nur die Hälfte oder das Drittel der Dauer des Engeelementes ausmacht. Bei den Affrikaten vom Typ *gy* (*dʲj'*) und *ty* (*t'ɣ'*) hingegen ist das Verschlusselement das ausgeprägtere und das Engeelement das reduziere. Die eigenartige Verbindung von Verschuß und Enge ist aber hier – wie wir es unter c) gesehen haben – ebenso charakteristisch wie bei anderen Typen. Dies wird übrigens auch durch die Tatsache bewiesen, daß wir das Engement nach der Sprengung des Verschlusses auch durch das Kehlkopfmikrophon gut verfolgen können: Abb. 33: *hattyú* (*hãt'ɣ'ũ*) «Schwan», Abb. 34: *itatja* (*itãt'ɣ'ã*) «er trinkt ihn». Bei diesen langen palatalisierten Affrikaten erreicht das Verhältnis des Verschlusses und der Enge sogar die Proportion 3:1, während sie bei den kurzen im allgemeinen 1:1 oder $1:\frac{3}{4}$ ist.



Außer der Zweimomentigkeit bestätigen derartige Analysen noch eine andere wichtige Tatsache: *Verschuß- und Engemente können voneinander abgegrenzt werden und folgen zeitlich nacheinander*, sie

werden also nicht gleichzeitig gebildet¹. Die Verkündung des Prinzips der Gleichzeitigkeit¹ widerspricht übrigens selbst den einfachsten Beobachtungen. Wie wäre es zum Beispiel möglich, das Element *t'* des ungarischen Wortes *hattyú* «Schwan» gleichzeitig mit dem spirantischen *χ'* zu bilden, wenn das Geräusch nur nach Sprengung des Verschlusses, nach Auflösung der großen Berührungsfläche entstehen kann?

e) Den zusammengesetzten Charakter der Artikulation der Affrikaten können wir übrigens auch an Aufnahmen, die mit dem Pegelschreiber von hoher Geschwindigkeit («High Speed Level Recorder») gemacht wurden, beobachten. Dieser moderne elektroakustische Apparat registriert die während der Bildung der Sprechlaute auftretenden Schalldruckschwankungen² und weist auf, daß bei der Bildung der Affrikaten der Schalldruck während des Verschlusses herabsinkt, und selbst nach dem Verschlußmoment nicht plötzlich in die Höhe steigt wie bei den echten Verschlußlauten, sondern sich zwischen dem Druckniveau des Verschlusses (0-Linie) und des folgenden Vokals bewegt (Abb. 35: *acél* [âtsël] «Stahl» und Abb. 36: *a cél* [â tsël] «das Ziel»). Die Schalldruckkurve der

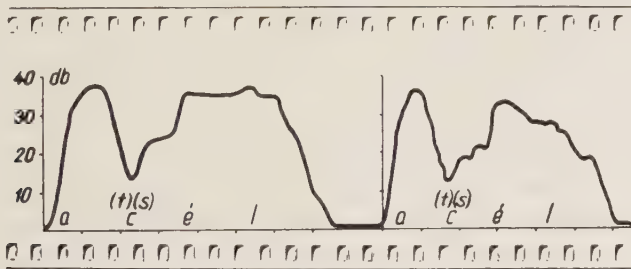


Abb. 35-36

Affrikaten zeigt also die Verbindung der charakteristischen Druckformen der Verschlußlaute und der Engelaute. Ähnliche Beobachtungen haben diesbezüglich auch *I. Mahnken* und *M. Braun* im Verlauf der Analysen von russischen Schalldruckaufnahmen gemacht: «Die Affrikata *c* (ts) bildet im allgemeinen zwei selbständige SDr-e über *t*- und über *s*-Element.» Unter dem Schalldruckbild 11 steht die Erklärung: «Auflösung des *c* in *t+s*», dann unter 12:

¹ Vgl. *Belgeri*, *Lazicius* und andere.

² *Hegedüs*, 1955, S. 181-214.

«Auflösung des \check{c} in $t + \check{s}'^1$.» Die beiden Momente der Schalldruckkurven der Affrikaten zeigen ferner I. Fónagys Aufnahmen².

Aus dem Obigen können wir also ersehen, daß die mit verschiedenen Forschungsmethoden erhaltenen Ergebnisse in vollem Einklang miteinander stehen und die Momente des Artikulationsverlaufes von verschiedenen Seiten her zeigen. Aus den vielseitigen Untersuchungen geht klar hervor, daß *die Affrikaten in artikulatorischer Hinsicht nicht einheitliche Laute sind, sondern einen zusammengesetzten Charakter haben*. Sie werden gekennzeichnet durch ein Verschlußelement und ein homorganes Engeelement. Ja, wir können sogar sagen, daß sie Verbindungen von Verschlußlauten und homorganen Englauten darstellen, die sich hinsichtlich der Stelle und der Weise ihrer Bildung einander anpassen, infolgedessen das eine oder das andere Element – eventuell beide – an Dauer, Muskelspannung, Stimmlippentätigkeit oder sonst etwas reduziert werden³.

5. Bei unseren Untersuchungen müssen wir auch in Betracht ziehen, welchen *akustischen Eindruck* die Affrikaten im allgemeinen in den Hörenden erwecken. Aus den bisherigen Beiträgen der Fachliteratur können wir feststellen, daß die diesbezüglichen Urteile sowohl der Phonetiker als auch der Laien von hochgradiger Unsicherheit und Schwankung zeugen. Einige nehmen die Affrikaten akustisch als die Verbindung zweier Elemente wahr, also als $t + s$, $t + \check{s}$, $d + z$, $d + \check{z}$, $t' + \check{x}'$, andere sprechen von «gemischter» Klangfarbe, andere wieder vom Eindruck einer «einheitlichen» Klangfarbe. Solche, die in ihrer Muttersprache keine Affrikate haben, beurteilen den akustischen Eindruck im allgemeinen als veränderlich, als nicht einheitlich und halten sie für die Verbindung zweier Laute. Solche aber, die selbst diese Laute gebrauchen, fassen sie oft als einheitliche Klangeindrücke auf.

Um den akustischen Eindruck zu analysieren, hat P. Passy⁴ das auf der Wachswalze des Phonographs aufgezeichnete spanische Wort *ocho* umgekehrt abgespielt und die Lautfolge *ošto* gehört. Auf

¹ Mahnken und Braun, S. 291.

² Fónagy, S. 402, Tafeln 19, 20 und 21.

³ Auch Meyer und Gombocz (Zur Phonetik der ungarischen Sprache. Upsala, 1909, S. 54) halten die Affrikate *c* (*ts*) für eine reduzierte Lautverbindung: «Man sieht, daß im allgemeinen die Lösung des *t*-verschlusses zur Überenge von einer Reduktion der Dauer begleitet ist», J. Forchhammer aber schon alle Affrikaten: «die Affrikaten können... als *reduzierte Lalenverbindungen* bezeichnet werden» (Z. f. Phon. 7, 405, 1953).

⁴ Passy, P., *Petite phonétique comparée*. Leipzig/Berlin 1922.

Grund dessen hat er festgestellt, daß die Affrikate *tʃ* akustisch nicht einheitlich ist, sondern genau aus zwei Elementen: *t* + *ʃ* besteht. Ähnliche Beobachtungen haben *I. Soames*¹ und *E. H. Tuttle*² bei englischen Affrikaten, *E. Richter*³ bei italienischen Affrikaten gemacht. *E. Richter* nahm die zu untersuchenden Wörter und Sätze nicht auf den Phonographen, sondern auf die Schallplatte auf, dann sie sechs Personen die umgekehrt abgespielte Aufnahme beobachten ließ und ihre Schalleindrücke aufzeichnen. Ihre prinzipielle Annahme war, daß beim umgekehrten Abspielen kein bekannter Wortkörper die Beobachter beeinflussen kann, da die aufeinanderfolgenden Laute ihnen unbekannte Wörter darstellen. Somit müßten die Affrikaten, wenn sie akustisch einheitliche Laute sind, auch umgekehrt gespielt einheitlichen Eindruck auf die Zuhörer machen. Die Protokolle über die subjektive Wahrnehmung waren sehr lehrreich, denn sie ergaben, daß die Beobachter die italienischen Affrikaten akustisch folgendermaßen wahrgenommen hatten: 1. als Verbindungen von zwei Lauten (sie notierten also z. B. *ʃt*) wurden sie in 140 Fällen (59,6%) aufgefaßt, 2. als Engelaute in 52 Fällen (22,1%), 3. als Verschußlaute in 39 Fällen (16,6%) und endlich 4. als Affrikaten einheitlicher Klangfarbe nur in 4 Fällen (1,7%)⁴. Es mutet uns sonderbar an, daß *E. Richter*, trotzdem die Beobachter in der Mehrzahl der Fälle die Affrikaten als Lautverbindungen beurteilten (beinahe 60%!), die Schlußfolgerung von ihren Untersuchungen zieht, daß sie «einheitliche Laute» sind. Dieses Beispiel beweist auch, wie nachteilig die Vorurteile die wissenschaftlichen Arbeiten beeinflussen können. Die Vermengung des akustischen Gesichtspunktes mit dem phonematischen hat mit Fug *J. Forchhammers* scharfe Kritik herausgefordert, wonach es dann offenbar wurde, daß die zwei Gesichtspunkte nicht identisch sind und einander auch nicht immer decken.

Die Methode des Umgekehrtspielens habe ich auch bei der Untersuchung des akustischen Eindrucks der ungarischen Affrikaten angewendet. Ich nahm Wörter, die lange und kurze Affrikaten wie auch die Folge von Verschußlaut und Engelaute enthielten, von 4 Versuchspersonen aufs Magnetophonband in Aussage- und Frageform auf, dann ließ ich 10 wissenschaftlich gebildete

¹ *Soames*, S. 38.

² *Tuttle*, S. 326.

³ *Richter*, S. 1–38.

⁴ *Forchhammer*, 1941, S. 13.

Linguisten das umgekehrt abgespielte Material abhören. Vom Zweck des Experiments hatten 7 Beobachter nichts gewußt, und so hielten sie die Wörter für arabisch oder aus einer ihnen unbekannten Sprache stammend. Beim Umgekehrtspielen wird natürlich auch die Reihenfolge der Artikulation umgetauscht, und deshalb ist besonders die Beobachtung des Stimmensatzes und -absatzes (weich, gehaucht usw.) sehr lehrreich. Was wir in der gewöhnlichen Rede gar nicht bemerken (z. B. die feine Behauchung nach Sprengung des Verschlusses: *fonnyad* [fɒ̃ˈãdʰ] «welkt»), tritt hier scharf hervor.

Die Beobachter hatten die Lautelemente folgender Wörter beim umgekehrten Abspielen zu kontrollieren: *cica* (tsitsã) «Kätzchen», *cápa* (tsãpã) «Hai», *acat* (ãtsãt) «Kratzdistel», *léc* (lẽts) «Latte», *Vác* (vãts) «Waizen» (Ortsname); *csáp* (tšãp) «Fühler», *csacsi* (tšãtsi) «Eselchen», «Eselfohlen», *rács* (rãtsš) «Gitter», *kacsa* (kãtsšã) «Ente»; *dzsem* (džem) «Jam», *findza* (findžã) «Tasse», *lándza* (lãndžã) «Lanze»; *brindza* «eine Art Schafkäse»; *gyár* (dʲj'ãr) «Fabrik», *gyep* (dʲj'ep) «Rasen», *ágyam* (ãdʲj'ãm) «mein Bett», *vágy* (vãdʲj') «Sehnsucht», *lágý* (lãdʲj') «weich»; *tyúk* (t'x'ũk) «Henne», *atyám* (ãt'x'ãm) «mein Vater», *bátyám* (bãt'x'ãm) «mein älterer Bruder», *Váty* (vãt'x') (Ortsname); *Ózd* (Ortsname), *ózdi* «von Ózd», *ezsdi* (eždi) «er fleht», *Pisti* «Steff» (Kurzform von Stefan), *áztat* (ãstãt) «er weicht ein»; *fonnyad* (fɒ̃ˈãd) «es welkt»; *bodza* «Holunder»; *maharadzsa* (mãhãrãdžšã) «Maharadscha»; *megbuggyan* (megbuđj'ãn) «es blubbert»; *lepottyán* (lepɒt'x'ãn) «er plumpst herunter»; *biccen* (bičsen) «er tut eine kleine wackelnde Bewegung».

Die Beobachter beurteilten zu 100 %, also ganz einstimmig, die Affrikaten als die Verbindungen von zwei Lauten. Sie nahmen an der Stelle des umgekehrt gespielten *c* (ts) – *szt* – (st –) Laute wahr (*cica* – *asztiszt*, *cápa* – *apászt*, *acat* – *taszta*, *Vác* – *sztáv* usw.; *biccen* – *neszt'ib*, *rács* – *stár*, *kacsa* – *hastak*), an der Stelle des *dz* ein *zd* (*brindza* – *azdnirb*, *bodza* – *azdob*, *hazdob*, *azdob*), an der Stelle des *dzs* (dž) ein *zsd* (žd) (*dzsem* – *mezsd*, *findza* – *azsdnif*, *lándza* – *azsdnál*, *maharadzsa* – *azsdaraham*, *hazsdaraham*, *azsd'araham*), an der Stelle des *ty* ein *x'* (bzw. *j*) *t* (*tyúk* – *kujt*, *lepottyán* – *najtx'opel*, *najx'topel*, *nastyopel*, *najx't'opel*), an der Stelle des *gy* (dʲj') ein *j'd* (*gyár* – *rãjd*, *gyep* – *djep*, *Djep*, *ágyam* – *majdá*, *vágy* – *jd* [n?] *áv*, *lágý* – *jdál*, *jnal*, *jnyál*; *megbuggyan* – *najx'dubgem*, *nazsdubgem*, *najx'd'ubgem*). Umgekehrt geben also die

akustischen Elemente genau dieselbe Zweigliedrigkeit wieder wie auch die Artikulation selbst. Es ist also zweifellos, daß wir die ungarischen Affrikaten in akustischer Hinsicht nicht einheitlich nennen können; es ist aber auch natürlich, daß der akustische Eindruck nicht bei allen Affrikaten gleich ist. Einen verhältnismäßig mehr verschmolzenen Klangeindruck machen auf die Zuhörer *c* (ts) und *cs* (tš) wie auch *ty* (t'x') und *gy* (d'j'). Der Grund hierfür liegt darin, daß bei ihrer Bildung das Verschluß- bzw. das Engeelement in hohem Grade reduziert ist, außerdem ist die tatsächliche Lautung akustisch hauptsächlich durch das zweite, d. h. das Geräuschelement vertreten. So gibt bei *c* (ts) und *cs* (tš) das Zischgeräusch, bei *ty* (t'x') das Reibegeräusch den typischen Lautungscharakter an. Anders verhält es sich akustisch bei den Affrikaten *dz* und *dzs* (dž). Die Klangfarbe dieser Laute empfindet man als am wenigsten einheitlich, ja die meisten Zuhörer nehmen sie entschieden als Lautverbindungen wahr. Und das hat wirklich seinen guten Grund. Bei der Bildung dieser Laute ist das Verschlußelement nicht stumm, sondern stimmhaft, und es stellt eine ganz andere Klangfarbe dar als das folgende stimmhafte Engeelement. Die Stimmhaftigkeit und die Zischgeräusche von hoher Frequenz beim zweiten Element machen diese Affrikaten fast zweigliedrig. Beim umgekehrten Abspielen wird diese Lautverbindung völlig wahrnehmbar. Außerdem habe ich mit einer Gegenprobe bestätigt, daß die Verbindung von zwei Lauten das Wesen dieser Affrikaten ausmacht. In den auf Magnetophonband aufgenommenen Wörtern *Ózd*, *ózdi* und *ezsdi* haben bei umgekehrtem Abspielen alle Beobachter *dz* bzw. *dzs* (dž) als Affrikaten wahrgenommen: *dzó*, *idzó*, *idzse*. Auf diese Weise ist es verständlich geworden, warum *dz* und *dzs* (dž) besonders viel Unsicherheit hervorrufen und warum sie auch den Fachleuten der Rechtschreibung in Ungarn so viel Kopferbrechen machen.

Die subjektive akustische Untersuchung der Affrikaten hat also ergeben, daß diese hinsichtlich der Klangfarbe nicht einheitlich sind und einen gewissen Spielraum zulassen. Bei stimmlosen Typen wird die Stummheit des Verschlusses von Zisch- und Reibegeräuschen beinahe unterdrückt, und darum wirken sie etwas einheitlicher (*c*, *cs*, *ty*). Die stimmhaften aber – besonders *dz* und *dzs* (dž) – erwecken entschieden den Eindruck der Zweigliedrigkeit in den Zuhörern. Das umgekehrte Abspielen der Affrikaten zeigt, daß sämtliche ungarischen Affrikaten in der Tat reduzierte Laut-

verbindungen sind, sie können aber einen einheitlicheren oder einen weniger einheitlichen Klangfarbeneindruck deshalb erwecken, weil die Geräuschelemente bei den Stimmlosen das stumme Verschlußelement unterdrücken können und auf diese Weise wahrnehmungspsychologisch das Übergewicht erlangen.

6. Den Phonemwert der Affrikaten pflegt man auf Grund des «Sprachgebrauchs», des Urteils der «Sprachgemeinschaft», des «Sprachbewußtseins» zu bestimmen. Die Meinung der Forscher ist auch hinsichtlich der phonematischen Wertung sehr geteilt. Einige fassen die Affrikaten als *Einzellaute*, andere als *zusammengesetzte Laute* bzw. *Lautverbindungen* auf. Bei uns schreiben sie ihnen im allgemeinen monophonematischen Wert zu. Wenn wir aber etwas näher auf ihre Analyse eingehen, so sehen wir, daß es gar keine einfache Aufgabe ist, dies zu entscheiden. Dies geht auch aus den folgenden Zeilen von *Laziczzius*¹ hervor: «Ob etwas ein einheitlicher Laut oder eine Lautverbindung ist, das ist – nach unserer Meinung – in erster Linie und hauptsächlich eine *sprachliche Frage*, die zu entscheiden nur mit Beachtung der in Frage kommenden sprachlichen Verhältnisse möglich ist. Die Tatsache selbst, daß die ungarische Sprache im Anlaut der Wörter die Affrikaten duldet, obwohl sie sonst alle Konsonantenverbindungen aufzulösen oder zu vereinfachen sucht, zeigt, daß im Ungarischen die Affrikaten unbedingt Einzellaute sind.» Es ist klar, daß die Bestimmung der «in Frage kommenden sprachlichen Verhältnisse» keine leichte Aufgabe ist. Selbst *Laziczzius* unternimmt nicht, diese zu bezeichnen, und spricht bloß von der Auflösung der Konsonantenhäufungen im Anlaut, was an sich kaum als überzeugender Beweis angenommen werden kann.

*Trubetzkoy*² schreibt über die einschlägige monophonematische Wertung, daß solche Lautverbindungen dann den Wert *eines* Phonems haben, wenn (I) ihre Elemente in der betreffenden Sprache nicht in zwei Silben auftreten können, wenn (II) sie durch einheitliche Artikulationsbewegungen hervorgebracht werden und wenn (III) überdies ihre Dauer die normale Dauer der Einzellaute nicht übersteigt.

Nach *Trubetzkoy's* Definition ist also der monophonematische Wert der ungarischen Affrikaten sehr unsicher, denn: (I) ihre

¹ *Laziczzius*, 1944, S. 80.

² *Trubetzkoy*, S. 50.

Elemente können manchmal in zwei Silben kommen¹, (II) ihre Artikulation hat zwei Momente und (III) endlich, ihre Dauer übersteigt den normalen Dauerwert der Einzellaute. Auf diese letztere Tatsache hatte ich schon während der ungarischen Affrikaten-debatte hingewiesen und auf Grund von Messungen erwiesen, daß die Dauer der Affrikaten im allgemeinen etwas länger ist als die der anderen Gruppen von Konsonanten. Übrigens haben auch *Meyer* und *Gombocz*² mit Messungen bestätigt, daß die Dauer der Affrikaten bedeutend länger ist als die der Verschlußlaute und Engelaute und manchmal sogar das Doppelte der der Einzellaute ausmacht: «Im besonderen ist die Gesamtdauer des *ts* in *táts* und *táts*, ca. 22. h. s., keineswegs gleich der Dauer eines einfachen Konsonanten, wie das nach *Balassa*... der Fall sein soll, sondern beträgt ungefähr das Doppelte derselben³.»

Seitdem haben neuere Untersuchungen und Dauermessungen diese These bestätigt. Die Durchschnittswerte der hier vorgeführten Affrikaten sind die folgenden: 1. *c* (*ts*) = 19,8 φ (*t*-Element = 6,17 φ , *s*-Element = 13,7 φ), 2. *cc* (\overline{ts}) = 30,7 φ (\overline{t} : 11,9 φ , \overline{s} : 18,8 φ), 3. *cs* (*tš*) = 16,2 φ (*t*: 5 φ , *š*: 11,2 φ), 4. *ccs* ($\overline{tš}$) = 26,7 φ (\overline{t} : 12,7 φ , \overline{s} : 14 φ), 5. *ty* (*t'x'*) = 18 φ (*t'*: 10,5 φ , *x'*: 7,5 φ), 6. *tty* (*t'x'*) = 30,8 φ ($\overline{t'}$: 23,1 φ , *x'*: 7,6 φ), 7. *dz* = 15,5 φ (*d*: 5 φ , *z*: 10,5 φ), 8. \overline{dz} = 25 φ (\overline{d} : 15 φ , \overline{z} : 10 φ), 9. *dzs* (*dž*) = 11,7 φ (*d*: 5 φ , *ž*: 6,7 φ), 10. *dzs* ($\overline{dž}$) = 19,8 φ (\overline{d} : 12,5 φ , $\overline{ž}$: 7,3 φ), 11. *gy* (*d'j'*) = 13 φ (*d'*: 8 φ , *j'*: 5 φ).

Nach meinen Dauermessungen, die ich im Dorfe Jágónak (im Komitat Baranya) mit Schallplatten vorgenommen habe⁴, beträgt zum Beispiel die Durchschnittsdauer des *t* 8 φ , des *d* 6 φ , des *š* 7 φ , des *s* 10 φ und des *z* 6 φ ; *gy* (*d'j'*) = 6,75 φ , *ty* (*t'x'*) = 7,5 φ , *cs* (*tš*) = 14,2 φ .

Die Daueranalyse der Sendung des ungarischen Rundfunks⁵ hat ergeben: *t* = 7,7 φ , *s* (*š*) = 8,6 φ , *sz* (*s*) = 10 φ ; *c* (*ts*) = 13,3 φ und *cs* (*tš*) = 12 φ . Auf Grund der Daueranalyse der Vorlesung «Költöznek a madarak» (Die Vögel ziehen fort) im ungarischen

¹ «A magyar helyesírás szabályai.» – Regeln der ungarischen Rechtschreibung. Budapest, 1954, § 321: *ed-zitek* «ihr härtet ihn ab», *madzag* «Bindfaden», *ped-zi* «er beißt an», *lopód-zik* «er schleicht»; *hod-zsa* «Hodscha», *maharad-zsa* «Maharadscha» usw.

² *Meyer* und *Gombocz*, S. 54–55.

³ a. a. O., S. 55.

⁴ *Hegedüs*, 1941, S. 10–15.

⁵ Leitartikel der Zeitung «Szabad Nép» vom 24. November 1954.

Rundfunk ist $t = 6 \varphi$, $s (\check{s}) = 10 \varphi$, $sz (s) = 10 \varphi$, $c (ts) = 13 \varphi$ und $cs (tš) = 16 \varphi$.

Chlumský und *Hála* haben für die tschechischen Affrikaten gefunden, daß zum Beispiel die durchschnittliche Dauer von c und $č$ länger ist als die des Verschußlautes t und dem Mittelwert von s und \check{s} nahe kommt oder ihm gleich ist¹. Nach *O. von Essen*² ist die Zeitdauer der deutschen Affrikaten auch länger als der Mittelwert anderer Lautgruppen.

Aus dem Gesagten geht also klar hervor, daß es im Grunde genommen nicht leicht ist, den Phonemwert der Affrikaten zu bestimmen. Es erwachsen viele Unsicherheiten daraus, daß ihre Elemente manchmal zu zwei Silben gehören können (bei den langen Varianten immer), daß sie zwei Artikulationsmomente aufweisen und daß sie hinsichtlich der Dauer eine eigenartige Stellung einnehmen.

D. Jones reiht in seinem Werk «The Phoneme: Its Nature and Use» die Affrikaten in eine besondere Kategorie ein. Er sagt, man könne drei Gruppen von Sprachlauten unterscheiden: a) einfache Laute («simple sounds»), wie zum Beispiel m , n , l , f , z und die Vokale, b) Gleitlaute («gliding sounds»), zum Beispiel w und j , und endlich c) zusammengesetzte Laute («compound sounds»), deren Hauptvertreter die Affrikaten wären. Das Hauptcharakteristikum dieser letzten Gruppe sei auch in artikulatorischer Hinsicht die Zusammengesetztheit: «An affricate has a 'stop' and a 'plosion' which terminates with friction reminiscent of the corresponding fricative consonant³.»

Unseres Erachtens ist ein großer Teil der ungarischen Affrikaten (c , cs , gy , ty) auch trotz ihrer physiologischen und akustischen Zusammengesetztheit als monophonematisch zu bewerten, die Stellung ihrer kleineren Gruppe (dz , dzs) ist aber überaus zweifelhaft. Bildung, akustischer Eindruck und Dauerverhältnisse dieser Laute liegen an der Grenzlinie; deshalb ergeben sie nicht nur in unserer Rechtschreibung Probleme, sie haben sogar im Laufe der Sprachgeschichte Unsicherheit hervorgerufen und Lautwandel zur Folge gehabt. Hierauf deutet die Tatsache, daß ihre Elemente als selbständige Laute auch Formen mit Lautumstellung erzeugt haben.

¹ *Hála*, S. 91.

² *v. Essen*, S. 94.

³ *Jones*, S. 4–5.

*L. Deme*¹ führt zahlreiche Beispiele an, aus denen er folgende Schlußfolgerung zieht: «Die Affrikate *dz*... tritt in gewissen Fällen mit Lautumstellung als *zd* auf, sie kann also kaum ein Laut sein.»

7. Die phonetische Bezeichnung der Affrikaten zeugt ebenfalls von großer Unsicherheit und auffallender Schwankung. *Szende Riedl*² (Mansuet) schließt sie aus der Gruppe der «einfachen» aus, er nennt sie «Mischkonsonanten», und dementsprechend meint er, daß *gy* durch die Buchstabenfolge *d+j (i)*, *ty* durch *t+j (i)*, *c* durch *t+sz*, *cs* durch *t+s* und endlich *ds* durch *d+s* wiedergegeben werden sollen. *Ĵ. Balassa* faßt die Affrikaten als Diphthonge auf und empfiehlt folgende Buchstaben zu ihrer Bezeichnung: *cs* = *ts*, *ds* = *dž* (a. a. O. 93–94), *c* = *t^s*, *dz* = *d^z*, *cs* = *t^š*, *dzs* = *dž^z*. *G. Bárczi*⁴ meint, einzelne Zeichen der alten finnisch-ugrischen Umschrift seien grundsatzmäßiger, also: *c*, *z*, *č*, *ž*. *E. Setälä*⁵ hat in der finnisch-ugrischen phonetischen Transkription folgendes: *c* = *t+s*, *č* = *t'+š'* oder *t+š*, *z* = *d+z* oder *d+s*. Im System der APhI dürfen Affrikaten mit zwei, aber auch mit einem Buchstaben bezeichnet werden: «Affricates may be represented either by single lettres or by digraphs; sometimes one method is more convenient and sometimes the other⁶.» *D. Jones* meint, zur Wiedergabe englischer oder deutscher Affrikaten seien Doppelbuchstaben zu empfehlen, zur Bezeichnung der Affrikaten türkischer und indischer Sprachen aber eher Einzelbuchstaben. *E. Richter*⁷ gibt die italienischen Affrikaten phonetisch mit zwei Buchstaben wieder und bezeichnet das an Dauer reduzierte Element mit einem kleineren Buchstaben: *č*. Ähnlich gebraucht *Ĵ. Forchhammer* zwei Buchstaben, und bei Affrikaten verbindet er die Buchstaben mit einem Bogen, während er bei gewöhnlichen Lautverbindungen den Bogen fortläßt (z. B. *čy* = Affrikate; *ts* = Verbindung zeitlich aufeinanderfolgender Konsonanten). Dagegen schreibt *B. Hála*: «Il est préférable de les marquer, dans la transcription phonétique, par des signes simples, en accord avec leur caractère d'unité articulatoire et acoustique⁸.» Und seine Affrikatenzeichen sind: *c*, *z*, *č*, *ž*⁹.

¹ *Deme*, S. 16.

² *Riedl*, Magyar Hangtan.

³ *Balassa*, S. 130.

⁴ *Bárczi*, S. 28.

⁵ *Setälä*, S. 34.

⁶ *Jones*, S. 5.

⁷ *Richter*, S. 23–38.

⁸ *Hála*, S. 93.

⁹ *Id.*: S. 94.

Da die phonetische Schrift die phonetische Natur der Sprachlaute widerspiegeln soll, sind wir der Meinung, daß wir die Affrikaten unbedingt mit zwei Buchstaben wiederzugeben haben. Dies allein kann die zwei Momente der Artikulation richtig widerspiegeln.

Zusammenfassung

Hinsichtlich der Beurteilung der Affrikaten herrscht im allgemeinen ein Durcheinander in der phonetischen Fachliteratur. Die einander widersprechenden Meinungen haben überall in der Welt manche Debatten entfesselt. Einige fassen die Affrikate als einen einzelnen Laut auf, andere wieder als die Verbindung zweier Laute. Einige Forscher teilen sie der Gruppe der Verschußlaute, andere der der Reibelaute, andere hingegen einer gesonderten Gruppe der zusammengesetzten Laute zu. Die Gesichtspunkte der physiologischen, akustischen und phonematischen Bewertung hält man im allgemeinen nicht auseinander, sondern man vermengt sie miteinander. Der Verfasser hat diese Gesichtspunkte einzeln hergenommen und auf dieser Grundlage die Momente des *Verlaufs der Artikulation* an Palatogrammen, Oszillogrammen, an mit automatischen Tonhöhenschreibern gemachten Bildern studiert; er hat an normal bzw. rückwärts gespielten Magnetophonaufnahmen den *auf die Zuhörer gemachten akustischen Eindruck* untersucht und endlich sich mit der *phonematischen Bewertung* und der *phonetischen Umschrift* der ungarischen Affrikaten befaßt.

1. Palatogramme ungarischer Affrikaten lassen erkennen, daß die Verschußbildung ein wichtiges Moment der Artikulation ist. Die Ausbreitung der Berührungsfläche von Zunge und Gaumen ist je nach Affrikatentypen und der effektiven Zeitdauer verschieden. Am kleinsten ist sie bei *dz* und *ts*, größer bei *dž* und *tš*, und am größten ist sie in der palatalisierten Gruppe (*d'j'*, *t'x'*).

2. Die von den Affrikaten gefertigten Kymogramme, Oszillogramme, Tonhöhenbilder und Schalldruckbilder zeigen, daß von einer gleichzeitigen Artikulation der Elemente der Affrikaten kaum die Rede sein kann. Zuerst entfaltet sich das Bild des Verschußelementes und erst dann das des Engeelementes, die beiden können voneinander abgegrenzt werden. Die Zeitdauerverhältnisse der Affrikatenelemente, die Dauerreduktion der Elemente sind je nach

den Typen verschieden. Bei den Typen *ts*, *dz*, *tʃ*, *dʒ* wird eher die Zeitdauer des Verschlusses (in der kurzen Form), in der palatalisierten Gruppe (*d'j'*, *d'ɲ'*) hingegen das spirantische Element reduziert. Aber es kann nicht nur die Zeitdauer der Elemente, sondern auch die Stimmlippentätigkeit, die Muskelspannung oder aber auch irgendein anderes Artikulationsmoment reduziert werden.

3. Auf Grund von umgekehrt gespielten Magnetophonaufnahmen haben die Zuhörer die Affrikaten ausnahmslos als Verbindungen von zwei Lauten beurteilt, während sie diese bei normalem Abspielen als Einzellaute beurteilten (Meinungsverschiedenheiten kamen nur bei den Affrikaten *dz* und *dʒ* vor).

Der Verfasser ist der Meinung, daß ein großer Teil der ungarischen Affrikaten trotz ihrer artikulatorischen und akustischen Zusammengesetztheit einen monophonematischen Wert zu haben scheint – obwohl die monophonematische Bewertung keine genauen und sicheren Kriterien hat –, da sie sich auch im Laufe der Sprachgeschichte als Einzellaute benommen haben. Nur die Stellung der Laute *dz* und *dʒ* beurteilt er als problematisch, da diese auch heute noch den Eindruck auf die Zuhörer machen, am meisten Lautverbindungen zu sein, und da deren Elemente im Verlaufe der ungarischen Sprachgeschichte oft umgetauscht wurden und dies darauf hindeutet, daß die Sprecher sie im allgemeinen als die Verbindung von zwei Lauten empfanden.

4. Zur phonetischen Bezeichnung der Affrikaten schlägt der Verfasser vor, auch weiterhin zwei Buchstaben zu verwenden, um den Charakter der Elemente auszudrücken. Diese Weise spiegelt nämlich am treuesten die Zweiteiligkeit des Verlaufs der Artikulation, die wirklichen Bildungsmomente wider. Ohne Zweifel soll die phonetische Bezeichnung nicht die phonematische Wertung, sondern die Feinheiten der Lautbildung wiedergeben.

Summary

The phonetic literature displays a general confusion concerning the evaluation of affricates. Contradictory opinions have set off many controversies all over the world. Some writers rate the affricate as a single sound, others as a combination of two sounds. Some allocate them to the stops, others to the fricatives, and still others to a separate category of compound sounds. As a rule the different levels of analysis – physiological, acoustic, phonemic – are not kept separate, but rather confused with each other. The author has taken these levels one by one and on this basis has studied the factors contributing to the articulation through palatograms, oscillograms and pitch-meter traces; he

has investigated the *acoustic impression on the hearer* by means of tape-recordings played both forwards and backwards, and lastly he has concerned himself with the *phonemic interpretation* and the *phonetic transcription* of the Hungarian affricates.

Résumé

En ce qui concerne les affriquées, la confusion règne dans la littérature phonétique technique. Les opinions les plus contradictoires ont déchainé bien des débats de par le monde. Certains considèrent les affriquées comme un son unique, d'autres comme la combinaison de deux sons. Certains chercheurs les rattachent au groupe des occlusives, d'autres à celui des fricatives, d'autres encore à un groupe séparé de sons complexes. Généralement, on ne sépare pas les points de vue physiologique, acoustique et phonématique, mais on les confond tous. L'auteur s'est placé à chacun de ces points de vue séparément et, sur ces bases, il a étudié les phases de l'articulation à l'aide de palatogrammes, d'oscillogrammes, de graphiques faits avec des enregistreurs de mélodie automatiques; il a étudié l'impression acoustique produite sur l'auditeur avec des bandes de magnétophone jouées normalement et à l'envers et finalement il s'est occupé de l'évaluation phonématique et de la transcription phonétique des affriquées hongroises.

Literaturverzeichnis

- Bakó, E.: Hangtani tanulmányok. Debrecen 1937.
 Id.: A magyar mássalhangzók rendszeréhez. Magyar Nyelv 38, 167–172, 1942a.
 Id.: Mégegyszer az affrikátákról. Magyar Nyelv 38, 357–358, 1942b.
 Balassa, J.: Magyar fonetika. Franklin-Társulat, Budapest 1904.
 Bárczi, G.: Fonetika. Tankönyvkiadó, Budapest 1951.
 Belgeri, L.: Les affriquées en italien et dans les autres principales langues européennes. Grenoble 1929.
 Csűrű, B.: Szamosháti palatogrammok. Nyelvtudományi Közlemények 50, 64–70, 1936.
 Deme, L.: A hangátvétel a magyarban. Magyar Nyelvtudományi Társaság Kiadványai 69, Budapest 1943.
 von Essen, O.: Allgemeine und angewandte Phonetik. Akademie-Verlag, Berlin 1953.
 Fónagy, I.: Über die Schallfülle der ungarischen Vokale. Acta linguist. Acad. scient. Hung. 4, 383–425, 1954.
 Forchhammer, J.: Zur Lösung des Affrikatenproblems. Arch. néerl. Phon. exp. 17, 9–20, 1941.
 Id.: Kern- und Wendepunkt der Sprechwissenschaft. Z. Phonetik 5, 23–36, 1951a.
 Id.: Allgemeine Sprechkunde (Laetik). Winter, Heidelberg 1951b.
 Id.: Zum Affrikatenproblem. Z. Phonetik 7, 403–408, 1953.
 Id.: Einteilung der Sprachlaute. Akustisch oder artikulatorisch? Stud. Linguist. 8, 34–53, 1954.
 Gemelli, A.: La strutturazione psicologica del linguaggio studiata mediante l'analisi elettro-acustica. Ex aedibus academicis in civitate Vaticana 1950.
 Gerhardt, D.: Zum Thesaurus der menschlichen Sprachlaute. Z. Phonetik 2, 82–108, 1948.
 Id.: Noch einmal die schriftdeutschen Affrikaten. Z. Phonetik 6, 57–76, 1952.
 Gombocz, Z.: Magyar palatogrammok. Nyelvtudományi Közlemények 38, 193–204, 1908.
 Hála, B.: Une contribution à l'éclaircissement de la nature phonétique des affriquées. Z. Phonetik 6, 77–93, 1952.

- Hartnagel, S.*: Az affrikáták természetéről. Magyar Nyelv 32, 319–321, 1936.
- Hegedüs, L.*: Az affrikáták kérdéséhez. Magyar Nyelv 32, 17–20, 1936a.
- Id.*: Fonetikai széljegyzetek. Magyar Nyelv 32, 189–191, 1936b.
- Id.*: Die Natur der ungarischen Affrikaten. Arch. néerl. Phon. exp. 15, 97–102, 1939.
- Id.*: Elektro-akusztikai berendezések a beszéd- és a nyelvjáráskutatás szolgálatában. Magyar Nyelvtudományi Társaság Kiadványai 62, Budapest 1941.
- Id.*: Palatogramm-mérés. Nyelvtudományi Közlemények 51, 64–73, 1941–1943.
- Id.*: Hangnyomásmérések. Nyelvtudományi Közlemények 56, 181–214, 1955.
- Horger, A.*: Mi az affrikáta? Magyar Nyelv 31, 210–218, 1935.
- Id.*: A vegyes zárhangok kérdéséhez. Magyar Nyelv 32, 100–103, 1936a.
- Id.*: A vegyes zárhangok kérdéséhez. Magyar Nyelv 32, 259–260, 1936b.
- Id.*: Mi az affrikáta? Magyar Nyelv 38, 285–286, 1942.
- Jones, D.*: The Phoneme: Its Nature and Use. Heffner, Cambridge 1950.
- Kelemen, J.*: Affrikáta-e a *tl*, *dl* hangkapcsolat? Magyar Nyelv 32, 119, 1936.
- Laziczius, Gy.*: Az affrikáta-vita. Magyar Nyelv 33, 76–82, 1937.
- Id.*: Fonetika. Királyi Magyar Egyetemi Nyomda, Budapest 1944.
- Madzsar, I.*: Az affrikáták és a szótaghatár. Magyar Nyelv 32, 103–106, 1936.
- Mahnken, I.*, und *M. Braun*: Zum «expiratorischen Akzent» im Russischen. Z. Phonetik 6, 285–314, 1952.
- Martinet, A.*: Occlusives and Affricates with Reference to some Problems of Romance Phonology. Word 5, 116–122, 1949.
- Meyer, E.*, und *Ž. Gombocz*: Zur Phonetik der ungarischen Sprache. Upsala 1909. Sonderdruck aus «Le Monde Oriental» 1907–1908.
- Richter, E.*: Die italienischen *č*- und *š*-Laute. Arch. néerl. Phon. exp. 16, 1–38, 1940.
- Riedl, Sz.*: Magyar Hangtan. Prag/Leipzig 1895.
- Rousselot, L'abbé*: Principes de Phonétique Expérimentale I. Didier, Paris 1924.
- Setälä, E. N.*: Über transskription der finnisch-ugrischen sprachen. Finn.-ugr. Forsch. 1, 15–52, 1901.
- Soames, I.*: Introduction to Phonetics. London 1891.
- Stetson, R. H.*, *C. V. Hudgins* and *E. R. Moses Jr.*: Palatograms change with Rates of Articulation. Arch. néerl. Phon. exp. 16, 52–61, 1940.
- Tarnóczy, T.*: Die akustische Struktur der stimmlosen Engelaute. Acta linguist. Acad. Scient. Hung. (Budapest) 4, 313–349, 1954.
- Trubetzkoy, N. S.*: Grundzüge der Phonologie. Travaux du Cercle linguistique de Prague 7, 1939.
- Tuttle, E. H.*: On English CH and J, and other Similar Sounds. Die neueren Sprachen. Z. neusprachl. Unterricht 11, 321–328, 1903–1904.
- Vass, B.*: Les articulations Hongroises. Paris, ohne Jahr.

The “th-Sounds” and Genetics

By ROBERT WALLACE THOMPSON, Hong Kong

Some time ago I published some observations²⁴ on the paragraphs dealing with *the genetic component of language* in *The Elements of Genetics*² by C. D. Darlington and K. Mather. I attempted to show, in very general terms, that they were wrong *de jure*, and, in rather more detail, that they were quite wrong *de facto**. For example, in one of the areas (Ireland) with a very high frequency (70-75%) of blood group O, they claimed that the voiced and voiceless dental fricatives are common**. It is indeed probable that the use of these sounds is restricted to the North East corner: and the popular belief that “th-sounds” are common in Ireland is partly due to attempts by writers (such as *Somerville* and *Ross*) to represent in conventional spelling sounds unknown to Standard English. In the same context, I also pointed out that soundshifts, apparently unrelated to any demonstrable alterations in the racial composition of the Greeks, had taken place in Greek. Finally, I showed that in the speech of Jamaicans^{22***} of African, European and mixed stock, phonetic preference of both the individual and the group depends on cultural rather than racial background.

* *Lancelot Hogben*, “Human Biology and Human Speech”, *British Journal of preventive and social Medicine* 10: 63-74 (1956), attacks their methodology from the biologist’s stand-point and reaches fundamentally the same conclusions as those of this paper.

** I am aware that *Darlington* and *Mather* are led indirectly to the conclusion that there is a biological factor in the development of language by what they believe to be the coincidence of their two maps. I do not, of course, believe they think blood group O has any *immediate* connection with the ability of individuals to utter *th*-sounds!

*** *Rens*, speaking of the Chinese of Surinam, says that “the second generation born and bred in Surinam, speaks both Dutch and NE normally, without a trace of foreign, that is, non-Surinam, accent”.

In this paper I shall first deal with other errors in a map (p. 263) which *Darlington* and *Mather* exhibit. Afterwards I shall consider their categorical statement (p. 361) that "the fricative dental has arisen only in the Eurasian region". I shall then try to set out briefly a more acceptable explanation of linguistic *substratum* in accordance with what I believe to be views held by many other modern students of language.

1. *Some Minor Factual Errors*

The following preliminary comments refer to the Map shown on page 363:

- (i) A ð-sound existed in earlier Finnish and exists in various Finnish dialects today^{19*};
- (ii) a ð-sound also existed in proto-Hungarian²³;
- (iii) dental fricatives occur in at least three Altaic¹³ languages – Balkar, Shors (ð) and Bashkir (spoken in the Bashkir Autonomous Republic) where θ is the medial allophone of phoneme /c/ and ð the final.

2. *θ/ð Outside Eurasia*

There is no doubt that the sounds θ and ð are comparatively rare among the languages of the world. Nevertheless, we have already seen that they are not quite so rare in the Eurasian region as *Darlington* and *Mather* indicate. Beyond this vast region, Africa¹⁵ is the one area where it is largely true (in their words) that "the fricative dental has arisen only in the Eurasian region". In that continent, the dental fricatives do not appear in general, save in the Semitic^{11**} languages or in regions under their influence. However, they have been recorded in some dialects¹⁰.

The dental fricatives also occur in the Melanesian language Lifu 15, p. 680, while ð occurs in the Melanesian languages Malu²¹ and Tangoa 15, p. 677. Aleut⁶, a language related to the Eskimo cluster, has both the voiced and the voiceless dental fricatives.

Though the presence or absence of these sounds in Amerindian languages is highly variable, we have good reason to believe that their blood group distributions were remarkably uniform before the

* I am grateful to Professor *Alan S. C. Ross* of Birmingham University for information about the dental fricatives in the Finno-Ugrian group and for a number of points on their situation outside the Eurasian land-mass. Prof. *Ross* is not responsible for my interpretation of the facts.

** In the Arabic of Syria and Egypt θ / ð are usually replaced by other consonants, but in Iraq they are more commonly retained.

coming of the Europeans¹⁷ *. δ is present in Yuma (¹⁵, p. 1185), Chippe-
wayan (¹⁵, p. 1180), Otomi (¹⁵, p. 1186) Osage⁴ (spoken in Pawhuska,
Okla.). Both are absent from the Carib⁷ group, Tonkawa (Texas)¹,
Quileute (N.W. Washington), Yuchi (Central Okla.), Zuni (W. New
Mexico), Cœur d'Alene (N. Idaho), Kiowa⁸ (spoken in 1918 at
Anadarko, Okla. but formerly in what is now W. Montana). We
may thus come to the conclusion that the concentration of group O
and an inborn facility for the pronunciation of the *th*-sounds re-
ceives no more compelling support from the study of American In-
dians than from the speech habits of other peoples. It would be
interesting to hear *Darlington* and *Mather's* views on *Lehmann's* \underline{s} –
 \underline{z} -sounds¹², which lie very close to θ – δ , as for instance in Guatuso
and Chilanga. Of \underline{s} – \underline{z} in Guatuso, *Lehmann* says (p. 380): "*Dem*
Guatuso eigentümlich sind gewisse gelispelte Laute, wie sie etwa in unserer
Kindersprache vorkommen." Of Chilanga he says (p. 700): "*Ein gelispelter*
Zischlaut ähnlich wie im Guatuso ist \underline{s} , stimmhaft \underline{z} ."

Though by no means common sounds in the languages of the
world, θ/δ do occur very widely, and, I believe, spontaneously in a
great many language groups and among many peoples; but there is
no evidence to show that individuals able to pronounce them – in
Europe or elsewhere – are mainly of blood group O.

3. Genetics and Linguistic Substratum

In short, the isogens for the O iso-agglutinin do not closely
coincide with isophons for the dental fricatives θ/δ in Europe. It
seems unlikely that they do so in the greater Eurasian land-mass;
and if the "th-sounds" do occur in aboriginal America – a predomi-
nantly O-continent – they do so sporadically. Ability to pronounce
them cannot be proved to be of genetic origin.

Let us now consider the question more generally, from the point
of view of *substratum*. By *substratum* one implies survival of lexical or
phonetic elements of a language in one which has superseded it. Such
a *substratum* emerges slowly or otherwise where a whole population,
at first bilingual, progressively gives up the old for the new. An entire
first generation will then carry over its current speech habits into the

* ... "the Indians of South and Central America who even now are almost entirely
of group O may have been exclusively so before the coming of the Europeans." For more
details on this subject, see Chap. XIII. Mr. *Douglas Taylor* from *Magua*, Dominica, tells
me that a "number of Amerindian tribes with 90% or more O group have either or
both the *th*-sounds, not only *qua* sounds, but as phonemes".

use of the new language, together with some words describing specific phenomena for which it caters inadequately. It is likely that the speakers will also modify the grammar of the new language. In their own lifetime they will speak a pidginized form of it.

To a greater or less extent, the phonetic system of the new language as later generations will speak it will retain features transmitted to the second generation by their mothers from whose mouths their children learn to utter its sounds. The formative role of the mother and of the old folk in this process has no necessary connexion with their genes. How drastic and how lasting its influence will prove must depend on circumstances incident to the transition. It may happen that native speakers of a newly imposed language may be warriors who marry the women-folk of the conquered speech community. Likewise, it may happen that the conquerors bring in their own women folk, and thus maintain continuity with the pre-existing standard of the new tongue.

In this way the raw African of the XVII or XVIIIth century handed down certain speech habits to the West Indians of today. In my earlier article, I indicated that the alleged inability of the Jamaican negro (and poor white) to pronounce the dental fricatives which he replaces by the corresponding voiced and voiceless alveolar *d* / *t* is attributable to cultural circumstances to which genetics has no relevance. He says *dis*, *dat*, *tink*, etc. (*this*, *that*, *think*, etc.) because his father (one should rather say *mother*) did so before him. The so-called *Africanisms* of his speech came down to him through generations of contact with such speech. There is no call to invoke genetic predisposition either to conserve or to reproduce them by a process of delayed action.

It has been shown that slavery in the Spanish and Portuguese colonies was more paternal, more human (but not necessarily more humane) than in the English or Dutch⁵. A deliberate attempt was made to meet the spiritual needs of the slaves and some fundamental education was given by the clergy. In addition, a more serious effort was made by the Spaniards and Portuguese to plant their New World territories with peasants of European stock. As a result, their subjects of African descent did not greatly overwhelm in numbers those who spoke a European language in conformity with native usage, as happened in the English, French and Dutch colonies of the West Indies. Today there is no Spanish *creole* dialect in the New World, where the negroes of Cuba, of the Dominican Republic and of Porto

Rico pronounce the voiced dental fricative δ every bit as well as their white or mulatto compatriots. The bozal or "raw" negro may have spoken a jargon. His children lost all the old speech habits in contact with both whites and blacks who spoke "normal" Spanish.

In the last fifty years a relevant phonetic process has been going on in Papiamentu, the creolized Portuguese of Curaçao²⁵. Here the medial consonants $b - d - g$ are gradually being replaced by their fricative variants $\beta - \delta - \gamma$. This gradual sound-shift coincides with contact with Spanish. The latter has conferred prestige in business and social intercourse for over a century, during which it has provided up to 85% of the contemporary vocabulary. This is part of a general process of decreolization, under the influence of Spanish, but we have no indications of any substantial genetic modification of the largely African population during the same period. Similar social forces impell the Hong Kong civil servant to straighten out his r -sounds, and the Yorkshire-born subaltern his o and u sounds. Such forces affect both individuals and communities.

Among linguists it was some years ago the fashion to speak of disappearance of certain sounds used by a community and their subsequent reappearance centuries later as the *deferred action of the substratum*, interpreted by some scholars in terms of the *profound action of hereditary tendencies* (A. Meillet, *Bull. Soc. Ling.*, Paris 1930)¹⁶. Even as late as 1934 *Tagliavini* maintained that there were *certain tendencies, particularly phonetic, which may be inherited and crop up again after many generations*. Like other eminent philologists such as *Wartburg*, *Dauzat* and *Meillet* himself, *Tagliavini* believed in the Lamarckian superstition. Such quasi-biological beliefs are therefore of no interest today except as an admission that the linguist once looked to genetics for an intelligible interpretation of at least one influence of the substratum.

Needless to say, the cogency of such an admission depends on whether it is conclusively possible to vindicate the existence of deferred action in the sense defined. The historic background of examples most commonly cited is misty. For instance, the consonant cluster $nd < nt$, $ng < nk$ of Osco-Umbrian allegedly turns up at a very late period in dialects of South Italy, and the u sound of modern French and German allegedly restores a previously suppressed Celtic vowel sound of the same region. We are on more solid ground, when we examine the highly characteristic shift $f > h$ in Spanish. This we commonly attribute to an *Iberian* substratum, though, until our own

century not attested in documents written before the XIVth. In the last fifty years the first reference has been pushed back by successive stages to the IXth century. Indeed Medieval research now leads us to believe that there is a complete *personal* link between those who substituted an *h*-sound for initial *f* in other languages of the pre-Romance* era and those who still do so today, or at least have done so until comparatively recent times. *Menéndez Pidal*, the father of modern Spanish philology, has shown that pre-Romance survivals in Spanish such as *perro*, *guijarro*, *barranco*, etc. are not documented before the XIIIth or XIVth century; but does this justify our belief in their spontaneous resurgence as part of the vocabulary of Castilian after many centuries of absence?

To be sure, we have no continuous documentation of the commonly unprintable four-letter words of Anglo-Saxon origin now found in modern, colloquial, sub-standard English; but no one on that account would seriously suggest that they are, in this century a mushroom growth in the Edwardian Dublin of *Joyce* or in the turbulent Anglo-Saxon world of *Hemingway*. In the same way it has become more and more apparent that certain sounds as well as certain words have been handed down from mother to child, over the centuries. In the words of *Menéndez Pidal* (¹⁶, p. 8):

Para suplir nuestra carencia de datos durante varios siglos no pensemos en una acción diferida, no en una fuerza latente inactiva cuya operación sólo reaparece muy tarde; pensemos sólo en un estado de actividad latente, un uso lingüístico propio del idioma socialmente deconsiderado, uso que primero se introduce subrepticiosamente en la lengua adoptiva como un bajo defecto, rechazado en absoluto por todo hablante culto, uso que después va ganando poco a poco el gusto de las generaciones sucesivas, a través de los siglos, en reacciones de llaneza y desafectación, llegando a ser tolerado entre la gente culta, hasta que al fin se hace preponderante y aun exclusivo.

*Morgan Watkin*¹⁷ falls into the same kind of error as *Darlington* when he attempts to revive the neglected theory of *Rhys* and *Morris Jones*. This theory seeks an explanation for some of the peculiarities of Welsh by postulating a non-Aryan population with a Hamitic language before the conquest of Wales by the Iron-Age Celtic invaders. To *Watkin* the existence of this suppositious Hamitic substratum is both a linguistic, and by the same token genetic, reality. Less cautious than *Darlington*, he enlists genetics alike in the service of

* This explanation is not accepted by Profs. *J. Orr* (Edinburgh) and *W. D. Elcock* (Westfield College) and others.

syntax, idiom (*sic*), accent and intonation, including the Welsh *hwyl* and the white North African Muezzin's call to prayer, though he does not mention the chant of the black North Nigerian. Linking Cader Idris with the Berber personal name Idriss is ingenious but scarcely transcends the level of special pleading. In fact, I see no need for an *ad hoc* explanation of pre-Celtic survivals in Welsh in the jargon of modern genetics or otherwise. Syntax and intonation are passed on from individual to individual, from generation to generation, as are vocabulary and sounds. The new language borrows them from the old in the stage bilingualism, as we have seen above.

Elsewhere (p. 29) *Morgan Watkin* associates the Basque blood groups and dental fricatives. There, his statement that the Basques have a strongly developed *th*-sound requires qualification. Basque z^{26} is not a dental fricative but a voiceless alveolo-dental *s* similar to the most usual *s* of Andalusia and Spanish America. I know of no Basque dialect where it could possibly be represented by *Watkin's* *tth* which I take to be *tθ* in my notation. The digraphs *ts* and *tz* represent affricates and not fricatives. Thus blood group O and dental fricative *θ* cannot be correlated in the Basque dialects since *θ* does not exist there. The dental fricative *ð* does, however, appear as a sub-phonemic medial variant of /d/ in some (but by no means all) dialects. *En passant*, I should like to point out that many Basques are bilingual and that, as a result, two of their dialects have been influenced by sounds in the neighbouring Romance languages, French and Spanish, which are the languages of prestige in the Basque country. *Souletin* uses the French *u*-sound (of *rue*, etc.) and *Guipuzcoan* the Spanish *jota* (*j* pronounced like the final consonant in Scottish *loch*). Earlier Basque seems uniformly to have used a *j*-sound similar to that of modern standard French *j* (a voiced palatal fricative). They pronounce it so today on the French side of the Pyrenees and in the Spanish province of Biscay, where it is losing ground to the *jota*. Here indeed, we see the bilingual process in action. *Morgan Watkin* should also have realized that the sound *ll* in Welsh (p. 29) is found in many languages, including English where it has, of course, only a *subphonemic* value, e.g. in English *cycling*, *Hitler*, etc. In English it is a member of phoneme /l/.

Before attempting to relate blood groups and alleged facility for utterance of certain sounds, I suggest that writers of his persuasion should acquire a deeper knowledge of the phonetic systems of the languages involved, and that they should get to know something

about the work that has been done by linguists in the last fifty years. As *Darlington* himself says in *The Facts of Life* (p. 422) ... "the great difficulty of merging two sciences lies in the obstinate refusal of the scientific specialist to understand his nearest neighbours."

Neither *Darlington* nor *Morgan Watkin* seem prepared to admit that the quality which other biologists (³, p. 133, 135) have called human educability is a far more essential component of language than the genetic basis of variability with respect to the organs of speech. Man has not responded to his environment in such a way that his organs of speech will force him (somewhat Jansenistically) every few generations to make sounds which his neighbours find difficult. His peculiar specific make-up has made him more and more flexible and less and less specialized as a sound-maker. That he is thus different from the parrot, cuckoo or jackass is indeed a consequence of his own genetic constitution.

Summary

Although the dental fricatives [ð], [θ] occur in many Western European Languages, we are not in a position to state that the areas where they are commonest are inhabited by communities with a high frequency of blood group *O*.

These sounds, though comparatively rare among the languages of the world, are by no means restricted to the Eurasian region.

It is also claimed that the phonetic preference of both the individual and the group depends on cultural rather than racial background.

Zusammenfassung

Wenn auch die Zahnreibelauten [ð], [θ] in vielen westeuropäischen Sprachen auftreten, sind wir doch nicht in der Lage, festzustellen, ob die Gebiete, in denen sie sehr verbreitet sind, von Menschen bewohnt werden, bei denen die Blutgruppe *O* besonders häufig ist.

Diese Laute sind – obwohl sie in den Sprachen der Welt verhältnismäßig selten vorkommen – keineswegs auf Europa beschränkt.

Sowohl der einzelne wie die Gruppen bevorzugen bei der Auswahl der Laute mehr kulturelle als rassische Gesichtspunkte.

Résumé

Bien que les fricatives dentales [ð] et [θ] se rencontrent dans de nombreuses langues d'Europe occidentale, nous ne sommes cependant pas en mesure d'établir si les territoires dans lesquels elles sont très répandues sont habités par des individus chez lesquels le groupe sanguin *O* est particulièrement fréquent.

Bien qu'ils apparaissent avec une relative rareté dans les langues du monde, ces sons ne sont en aucune façon restreints à l'Europe.

Dans le choix des sons, l'individu aussi bien que le groupe fait passer au premier plan des considérations culturelles bien plus que raciales.

Bibliography

1. *Boas, F.*: Handbook of American Indian Languages (Smithsonian Inst., Hamburg/ New York 1933).
2. *Darlington, C. D.* and *Mather, K.*: Elements of Genetics (Allen & Unwin, London 1949).
3. *Dunn, L. C.* and *Dobzhansky, Th.*: Heredity, Race and Society (The New American Library, New York 1954).
4. *la Flesche, Francis*: A Dictionary of the Osage Language (Bulletin of American Ethnology 109, Washington 1932).
5. *Freyre, G.*: Case grande e senzala; 4. ed. (J. Olympio, Rio de Janeiro 1943), cf. *Putnam, S.*
6. *Geoghegan, R. H.*: The Aleut Language (Smithsonian Inst., Washington 1944).
7. *De Goeje, C. H.*: Etudes linguistiques caraïbes (1909).
8. *Harrington, J. P.*: Vocabulary of the Kiowa Language (Bulletin of American Ethnology 84, Washington 1928).
9. *Hogben, L.*: Human Biology and Human Speech. Brit. J. prevent. soc. Med. 10: 63-74 (1956).
10. *Homburger, L.*: The Negro-African Languages, p. 42 (Kegan, London 1947).
11. *De Lacy O'Leary, E.*: Colloquial Arabic; 3rd. ed. (Kegan, Trench, Trubner & Co., London 1925).
12. *Lehmann, W.*: Die Sprachen Zentral-Amerikas; in Zentral-Amerika (Berlin 1920).
13. *Mathews, W. K.*: Languages of the U.S.S.R. (University Press, Cambridge 1951).
14. *Meillet, A.*: Bull. Soc. Ling. (1930).
15. *Meillet, A.* et *Cohen, M.*: Les Langues du monde, p. 854, Nouv. Ed. (Société de Linguistique de Paris, Paris 1952).
16. *Menéndez Pidal, R.*: Modo de obrar el substrato lingüístico. RFE t. XXXIV Cuadernos. 1-4.
17. *Mourant, A. E.* and *Morgan Watkin, I.*: Blood Groups, Anthropology and Languages in Wales and the Western Countries. Heredity 6: 13 (1952).
18. *Mourant, A. E.*: The Distribution of the Human Blood Groups (Blackwell, Oxford 1954).
19. *Ojansuu, H.*: Suomen lounai smurteiden äännehistoria XXX, II, p. 45 ff.
20. *Putnam, S.*: The Masters and the Slaves. Translation of: *Freyre, G.* cf. (Knopf, New York 1946).
21. *Ray, S. H.*: A comparative Study of the Melanesian Island Languages, p. 498 (University Press, Cambridge 1926).
22. *Rens, L. L. E.*: The Historical and Social Background of Surinam Negro-English (Amsterdam 1953).
23. *Szinnyei, J.*: Magyar nyelvhasználat; 7. Aufl., p. 36.
24. *Thompson, R. W.*: Nature 174: 10095 (1954).
25. *Tomás, T. N.*: Observaciones sobre el papiamento. NRFH 7: 183-189 (1953).
26. *Tovar, A.*: La lengua Vasca, p. 38 (Biblioteca Vascongada de los Amigos del País, San Sebastian 1950).

Author's address : Robert Wallace Thompson, University of Hong Kong, Department of Modern Languages, Hong Kong

Einige Ergebnisse von Untersuchungen der akustischen Struktur der Vokale

Von RAOUL HUSSON (Sorbonne)

Untersuchungsinstrumentarium

Wir haben uns der sogenannten Differenzial-Frequenzverstärker von Dr.-Ing. *Pimonow* systematisch bedient. Diese Vorrichtungen ermöglichen, innerhalb eines komplexen (periodischen oder nichtperiodischen) Gebietes die Intensität jedes beliebigen Frequenzbereichs nach Wahl zu erhöhen oder herabzusetzen. Jeder Apparat trägt am Anfang ein empfindliches Mikrophon und am Ausgang entweder einen Lautsprecher oder Kopfhörer. Wir haben Versuchspersonen beiderlei Geschlechts untersucht mit ausgebildeten und mit nichtausgebildeten Stimmen, vorwiegend Künstler mit kräftiger Stimme in allen Registern und in ihrem ganzen Stimmumfang. Die Kontrolle der Abänderungen der Teilton-Zusammensetzungen läßt sich leicht am Leuchtschirm eines Frequenzanalysators ablesen. Die Abänderungen der Spannung des Sphincter laryngis lassen sich ohne Schwierigkeit laryngo-stroboskopisch beobachten. Die Abänderungen im Bucco-Pharynx und am Velum können am radioskopischen Schirm betrachtet werden.

Nicht-Nasalvokale

*G. Beckmann*¹ (Kiel) konnte 1956 nach Entfernung des bucco-pharyngalen Ansatzrohrs die akustische Struktur des primären Glottistons experimentell erforschen. Dieser primäre Glottiston ist ein Sirenenton, d. h. er besteht aus einer Mischung von merklich gleich starken Frequenzen, von denen die tiefste die der Schwingungen der Stimmlippen ist und deren höchste mit der Spannung

der Glottisschläge (2000 Hz bei schwachen Schlägen, etwa 5500 Hz bei sehr starken Stimmen) zunimmt (Abb. 1, senkrechte Schraffierungen).

Unsere experimentelle Untersuchung hat gezeigt, daß die Ausbreitung im Bereich des bucco-pharyngalen Ansatzrohrs, und zwar von der Glottis bis zu den Lippen, die oben beschriebene Teiltonzusammensetzung folgendermaßen ändert:

1. Die Einschnittfrequenz des bucco-pharyngalen Gebietes – laut einem von *Yves Rocard*² eingeführten wichtigen Begriff – ist grob zwischen 2200 und 2500 Hz enthalten. Unter dieser Frequenz wird die Luft dieses Raums «in Bausch und Bogen» zurückgeschleudert, und es können an verschiedenen Stellen des Raums Resonanzen «hängen bleiben». Oberhalb dieser Frequenz gibt es eigentlich eine «Ausbreitung»; die sehr verschiedenen Resonanzbedingungen werden nicht mehr verwirklicht, und die Teile der Teilton-Zusammensetzung werden ohne nennenswerte Veränderungen übertragen³.

2. Hieraus ergibt sich, daß, wenn die «Spitzen» des endgültigen Vokalspektrums tiefer sind als die Einschnittfrequenz, sie tatsächlich Resonanzräumen entsprechen, die von der Gestalt des Raums bestimmt werden. Liegen sie höher als die Einschnittfrequenz, so entsprechen sie einfach einigen starken Teiltönen der Teiltonzusammensetzung, wenig verändert auf Grund des Durchgangs durch den Raum. Folglich muß die Bezeichnung «Vokalformant» ausschließlich den «Intensitätsgipfeln», die tiefer sind als die Einschnittfrequenz, zugeordnet werden, da nur sie von dem Einfluß des Resonanzraums auf die Teilton-Zusammensetzung der Glottisabhängig sind.

3. Wir haben festgestellt, daß jeder menschliche Vokal aus zwei tiefer als die Einschnittfrequenz liegenden Hauptformanten besteht, die ihm die eigentliche Vokalklangfarbe (oder seine Vokalität) ausdrücken. Der stärkere Formant wird durch eine Mundresonanz bestimmt, und er braucht nicht unbedingt der höchste zu sein. Der andere entspricht einer im Pharynx entwickelten Resonanz. Beide sind eng verbunden (gekoppelt), wobei die Verbindung bei den «geschlossenen» größer ist als bei den «offenen» Tönen. Für jede Grundvokalität (oder Vokaltypus) hängen die Tonalbereiche, in denen sich jeder dieser beiden Formanten (oder Formantengebiete) befindet, nicht von dem Grundton ab, sie sind aber umfangreich (250–600 Hz).

4. Beim Verlassen des Mundraums ist also das Spektrum in sieben Tonalgebiete zerlegbar: die Formantengebiete F_1 und F_2 ;

die Vor-, Zwischen- und Oberformantengebiete A, B und C, die tiefer als die Einschnittfrequenz liegen; die Oberformantengebiete D und E, die höher als die Einschnittfrequenz liegen, getrennt etwa um 3500 Hz (Abb. 1). Wir haben feststellen können³, daß die akustischen Bestandteile jedes Gebietes folgendermaßen zur Klangfarbe jedes Vokals beitragen: a) die Gebiete F_1 und F_2 verleihen die Vokalität (den Vokalcharakter), da der «Mundgipfel» stets am wirksamsten ist; b) die Gebiete A, B und C (die tiefer liegen als die Einschnittfrequenz) verleihen die «Rundung» (oder die «Dicke») des Vokals; c) die Gebiete D und E (die höher sind als die Einschnittfrequenz) verleihen dem Vokal seinen «Glanz» (oder das «Durchdringende»), diese Aufgabe kommt vorwiegend dem Gebiet D zu, wogegen Gebiet E (das oft nicht vorhanden ist) dem Klang eine besondere «Härte» verleiht.

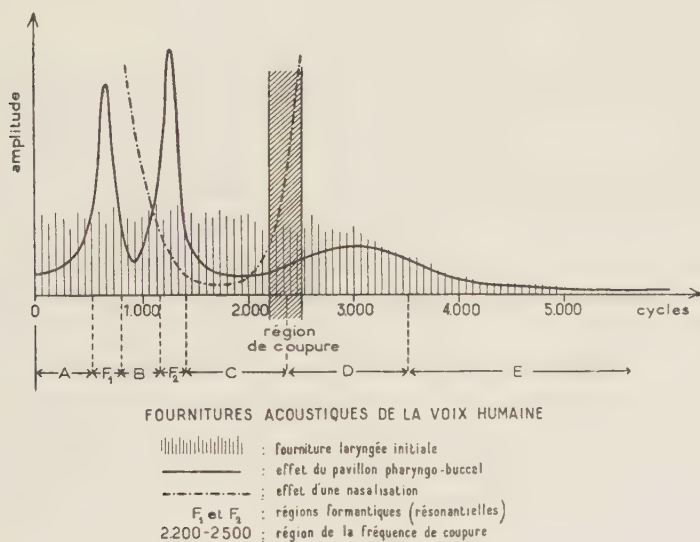


Fig. 1

5. Das gleichzeitige «Abkämmen» dieser beiden Formanten F_1 und F_2 auf jedem Vokal läßt eine von jeder Vokalität entblößte und ziemlich gleichförmige Klangfarbe fortbestehen. Sie hängt nur sehr wenig von dem erzeugten Vokal und von dem Grundton ab. Sie scheint aber für jede Versuchsperson kennzeichnend zu sein. Aus diesem Grund nennen wir sie «individuelle außervokalische Grundklangfarbe»³.

Nasalvokale

Die vier Nasalvokale (\tilde{a} , $\tilde{\varepsilon}$, \tilde{o} , $\tilde{\omega}$) unterscheiden sich von den vier entsprechenden offenen Vokalen a , ε , o , ω nur durch Überlagerung eines mit «Nasalisation» bezeichneten Vorgangs, der erzeugt wird durch eine mäßige und unvollständige Senkung des Velums während ihres Zustandekommens. Bis auf den heutigen Tag hatten Phonetiker und Akustiker geglaubt, daß die *Nasalisation* durch das Hinzukommen von «Sonderformanten», die von Resonanzfaktoren im Nasenraum abhingen, entstand. Aber alle Versuche, diese Faktoren zu bestimmen, blieben stets erfolglos. Mehr als 20 vermeintliche Formanten (zwischen 200 und 7500 Hz) wurden von den verschiedenen Forschern herbeigeholt, die aber eine lächerlich geringe Intensität aufwiesen. Auch die Versuche einer Synthese von Nasallauten durch Hinzufügen von diesen vermeintlichen «Nasalformanten» blieben stets völlig negativ.

In Wirklichkeit muß die Rolle des Nasenraums bei der *Nasalisation* der Vokale wie folgt analysiert werden: Der Rhino-Pharynx spielt die Rolle eines auf das bucco-pharyngale Gebiet aufgesetzten Filters «Passe-Bas», aber das Durchlaufen durch die Gänge und die Nasenlöcher setzt die Durchlässigkeit des Filters für sehr hohe Frequenzen herab und verwandelt den Filter in einen Filter «Enlève-Bande»⁴.

Diese Analyse erklärt die früheren von *Barczinski* und *Thienhaus*⁵ erzielten Ergebnisse, die damals für unerklärlich und irreführend gehalten wurden. Unsere Experimente haben ohne Schwierigkeit bestätigt, daß die *Nasalisation* bei jedem Vokal zustande kommt, wenn man die Intensitäten zwischen 1200 und 2000 Hz genügend ausmerzt (Abb. 1, gestrichelte Linie). Bei Versuchspersonen mit weitem Rhino-Pharynx senkt sich dieser Bereich bis 1000–1800 Hz, dagegen steigt er bis 1400–2200 Hz bei engem Rhino-Pharynx, wie wir durch Quer-Radiographien haben kontrollieren können. Umgekehrt unterdrückt man die *Nasalisation* oder setzt sie wenigstens entsprechend der künstlichen Verstärkung weiter herab, wenn man diesen Frequenzumfang vergrößert und ihn auf einen Nasalvokal einwirken läßt⁴.

Folgende Nebenerscheinungen begleiten noch die *Nasalisation*: 1. Gleiten der bucco-pharyngalen Formanten durch Veränderung der Koppelung im Bucco-Pharynx; 2. Behindernde Einschränkung der Glottistätigkeit; 3. Erscheinen von supervelaren Abzugstönen mit Hochfrequenzen, die im allgemeinen nicht stark sind⁴.

Zusammenfassung

Die Einschnittfrequenz des bucco-pharyngealen Gebietes ist grob zwischen 2200 und 2500 Hz enthalten. – Nicht-Nasalvokale: Die zwei «Hauptspitzen» des endgültigen Vokalspektrums, tiefer als die Einschnittfrequenz, entsprechen tatsächlich einer im Pharynx und im Mund entwickelten Resonanz. Die «Spitzen», höher als die Einschnittfrequenz, entsprechen einfach einigen starken Teiltönen der Teiltonzusammensetzung, die wenig verändert sind auf Grund des Durchgangs durch den Raum. Die zwei Formantengebiete verleihen die Vokalität. Die anderen Gebiete, die tiefer liegen als die Einschnittfrequenz, verleihen die «Rundung» des Vokals. Die Gebiete, die höher sind als die Einschnittfrequenz, verleihen dem Vokal seinen «Glanz». Das gleichzeitige Abkämpfen der beiden Formanten läßt eine von jeder Vokalität entblößte Klangfarbe fortbestehen, aber für jede Versuchsperson kennzeichnend, und wird «individuelle außervokalische Grundklangfarbe» genannt. – Nasalvokale: Die Rolle des Nasenraums bei der Nasalisation der Vokale muß wie folgt analysiert werden: Der Rhino-Pharynx spielt die Rolle eines Filters «Passe-Bas»; aber das Durchlaufen durch die Gänge und die Nasenlöcher verwandelt den Filter in einen «Enlève-Bande». Die Nasalisation kommt bei jedem offenen Vokal zustande, wenn man die Intensitäten zwischen 1200 und 2000 Hz genügend ausmerzt. Umgekehrt unterdrückt man die Nasalisation, wenn man diesen Frequenzumfang vergrößert und ihn auf einen Nasalvokal einwirken läßt. Gewisse Nebenerscheinungen begleiten die Nasalisation: Veränderung der Koppelung im Bucco-Pharynx, Gleiten der bucco-pharyngealen Formanten, behindernde Einschränkung der Glottistätigkeit, Erscheinen von supervalaren Abzugstönen von hohen Frequenzen (nicht stark).

Summary

The cutoff frequency of the bucco-pharyngeal region is roughly between 2200 and 2500 cps. – Non-nasalized vowels: The two peaks of the eventual vowel spectrum which are lower than the cutoff frequency correspond in fact to resonances developed in the pharynx and in the mouth. The peaks which are higher than the cutoff frequency correspond simply to certain strong components of the harmonic structure which are little altered by their transmission through the cavity. The two formant-areas are responsible for vocality. The other areas, when they are lower in frequency than the cut-off, give "rounding" to the vowel, and when higher than the cut-off they give the vowel its "brilliance". If both formants are simultaneously filtered out, a tone-colour persists which has been stripped of all vocality but is characteristic of each particular

speaker and is therefore termed "individual extra-vocalic basis tone-colour". Nasal vowels: the role of the nasal cavity in the nasalization of vowels must be analysed as follows: the naso-pharynx plays the part of a low-pass filter, but the passage of air through the nasal cavity and nostrils changes the filter into a band-stop. Nasalization arises with every open vowel if the energy between 1200 and 2000 cps are sufficiently eliminated. Conversely, nasalization is suppressed if this frequency range is amplified and applied to a nasal vowel. There are certain side effects accompanying nasalization: alteration of the coupling in the oro-pharynx, gliding of the bucco-pharyngeal formants, restriction of glottal activity, and the appearance of super-velar difference tones from high frequencies (not strong).

Résumé

La fréquence de coupure du pavillon pharyngo-buccal est grosso modo comprise 2200 et 2500 cycles. – Voyelles non-nasalisées: Les deux «clochers» principaux du spectre vocal définitif, inférieurs à la fréquence de coupure, correspondent à des états résonantiels développés dans le pharynx et dans la bouche. Les «clochers» supérieurs à la fréquence de coupure correspondent simplement à des partiels intenses dans la fourniture laryngée et peu modifiés par la traversée du pavillon. Les deux régions formantiques confèrent la vocalité. Les régions autres et inférieures à la fréquence de coupure donnent la plénitude de la voyelle. Les régions supérieures à la fréquence de coupure confèrent à la voyelle son éclat. L'écrêtement simultané des deux formants laisse subsister un timbre dépourvu de vocalité, mais caractéristique de chaque sujet, dit «timbre extra-vocalique individuel fondamental». – Voyelles nasalisées: Le rôle des cavités nasales dans la nasalisation des voyelles doit être analysé comme suit: le rhino-pharynx joue le rôle d'un filtre «passe-bas»; mais le passage par les méats et les narines le transforme en filtre «enlève-bande». La nasalisation apparaît sur toute voyelle ouverte lorsqu'on ampute suffisamment les intensités comprises entre 1200 et 2000 cycles environ. Inversement, l'amplification de cette bande de fréquences diminue ou supprime la nasalisation. Certains phénomènes secondaires accompagnent la nasalisation: altération du couplage pharyngo-buccal, glissement des formants, modification impédancielle de la fourniture glottique, apparition de sons d'écoulement sus-vélaires de haute-fréquences (peu intenses).

Literaturverzeichnis

1. Beckmann, G.: Arch. Ohr.-Nas.-Kehlk.-Heilk. 169, 196–201, 1956.
2. Rocard, Y.: Actualités Scientifiques et Industrielles. Vol. 222, S. 13–39. – Ibid.: Dynamique Générale des Vibrations. 2^e édit., S. 374 ff. Masson, Paris 1953.
3. Husson, R. et L. Pimonow: C. R. Acad. Sci. Paris 244, 1261, 1957.
4. Husson, R.: C. R. Acad. Sci. Paris 244, 2551, 1957.
5. Barczinski, L. und E. Thienhaus: Arch. néerl. Phon. exp. 11, 47–68, 1935.

Adresse des Auteurs: Prof. Raoul Husson, 10, rue Pierre-Picard, Paris 18 (France)

Regeln der deutschen Silbenmelodie*

Von ADALBERT MAACK, Braunschweig

Einleitung

In einer Reihe früherer Untersuchungen¹ habe ich mich mit den Formen der Lautmelodie beschäftigt. Ich glaube dabei den Nachweis geführt zu haben, daß der weitaus größte Teil (im Durchschnitt über 80%) aller Laute – ziemlich gleich, ob Vokale, Diphthonge oder stimmhafte Konsonanten – einen im wesentlichen geradlinigen Verlauf zeigt und daß die restlichen Laute in überwiegendem Maße eine nur *einfach* gekrümmte Melodieform aufweisen. Für die Silbenmelodie liegen m.W. ähnliche eingehende Untersuchungen noch nicht vor. *Wolfgang Bethge* hat kürzlich² das Thema im Zusammenhang mit den Abhöreraussagen über die Lautmelodie kurz gestreift, aber nur für das Verhältnis zwischen Vokal und *folgendem* Konsonanten und unter einem anderen Gesichtspunkt³. *Bethge* hat dabei festgestellt, daß sich die Abhörer bei der Beurteilung der Lautmelodie durch den *Silbenmelodie*verlauf beeinflussen lassen. Es erhebt sich darauf die Frage, ob man die objektive, gemessene Silbenmelodie, ähnlich wie die Lautmelodie, im wesentlichen durch eine Gerade bzw. eine höchstens *einfach* gekrümmte Kurve darstellen kann, oder ob die Silbenmelodie andere – gegebenenfalls welche – Formen aufweist, und welche Gesetze überhaupt für die Silbenmelodie gelten.

Die vorliegende Untersuchung stützt sich auf zwei phono-

* Anmerkung der Redaktion: Diese Arbeit gehört inhaltlich zu den Arbeiten desselben Verfassers, die im 1. Jahrgang dieser Zeitschrift veröffentlicht wurden – siehe dort S. 206 ff.; S. 216 ff. und S. 230 ff. –; aus drucktechnischen Gründen konnte sie erst hier veröffentlicht werden.

¹ *A. Maack*, 1938, S. 145 ff.; 1939, S. 27 ff.; 1955, S. 294 ff.

² *W. Bethge*, 1953, S. 320 ff.

³ Vgl. auch weiter unten.

metrische Textlisten: eine Erzählung, gesprochen von einem in Berlin wohnenden Arbeiter schlesischer Herkunft⁴, und einen Vorlesetext (aus «Gösta Berling» von Selma Lagerlöf), gesprochen von einem Wissenschaftler mit leichtem schlesischen Akzent⁵.

Zunächst ist die Frage zu klären, welche Silben zur Untersuchung heranzuziehen sind. Es gibt Silben mit einem, zwei und mehr Sonorlauten⁶. Die sehr häufigen Silben mit nur einem Sonorlaut, d. h. einem Sonanten als Silbenträger und einem oder mehreren stimmlosen Konsonanten, fallen natürlich von vornherein aus. In der Hauptsache müssen wir uns stützen auf zweisonorige Silben: mit einem Sonanten und einem (vorhergehenden oder folgenden) stimmhaften Konsonanten – wobei noch andere, stimmlose Konsonanten vorhergehen oder folgen können. Dreisonorige Silben sind nicht so häufig und kommen fast nur in der Reihenfolge: Konsonant–Sonant–Konsonant⁷ vor. Denn die Lautfolgen: Konsonant–Konsonant–Sonant und Sonant–Konsonant–Konsonant sind einerseits an sich noch seltener, andererseits sind dort, wo sie vorkommen, oft nicht alle Sonoren deutlich periodisierbar, d. h. der Melodieverlauf ist nicht eindeutig festzustellen. Dasselbe gilt natürlich in noch höherem Maße für vier- und mehrsonorige Silben. Wir sind deshalb für die Untersuchung auf die zweisonorigen Silben und die dreisonorigen in der genannten Lautfolge angewiesen.

Wenn nun die Silbenmelodie – ähnlich wie die Lautmelodie – einen überwiegend *geradlinigen* Verlauf zeigen würde, dürften zwischen den Melodiewinkeln⁸ der einzelnen stimmhaften Laute einer Silbe keine allzu großen Differenzen vorkommen, d. h. der Melodiewinkel des silbentragenden Sonanten müßte mit dem des vorhergehenden bzw. folgenden Konsonanten große Ähnlichkeit haben, m. a. W.: es müßte zwischen dem Melodiewinkel des Sonanten und dem des Konsonanten eine ziemlich hohe positive Korrelation bestehen. Man kann im allgemeinen nur zwischen *zwei* Variablen Korrelationen aufstellen: in unserm Falle also in zweisonorigen Silben zwischen dem Sonantenwinkel und dem Winkel des vorhergehenden bzw. folgenden Konsonanten. Dreisonorige Silben *als*

⁴ Phonom. Forsch., Reihe B, Bd. 3.

⁵ Phonom. Forsch., Reihe B, Bd. 6 (noch nicht erschienen).

⁶ Unter «Sonoren» seien in dieser Arbeit sämtliche stimmhaften Laute verstanden, also Sonanten (d. h. Vokale und Diphthonge) und stimmhafte Konsonanten.

⁷ Der Einfachheit halber sei im folgenden, wo kein Mißverständnis möglich ist, der Zusatz «stimmhaft» vor «Konsonant» fortgelassen.

⁸ Über die Berechnung des Melodiewinkels vgl. E. Zwirner und K. Zwirner, S. 4 ff.

Ganzes können nicht ohne weiteres korrelationsstatistisch untersucht werden⁹. Wir brauchen auf diese Silben jedoch nicht zu verzichten, wenn wir sie auf einfache Weise in die Lautfolgen: Sonant–Konsonant und Konsonant–Sonant *aufteilen*^{9a}.

Wir werden also zunächst die Korrelationen zwischen dem Sonanten- und dem Konsonantenwinkel für *alle* Lautfolgen Sonant–Konsonant bzw. Konsonant–Sonant innerhalb einer Silbe aufstellen, ohne Rücksicht darauf, ob die Laute in zwei- oder dreisonorigen Silben stehen. Als *Ganzes* werden wir die dreisonorigen Silben im zweiten Teil dieser Arbeit im Zusammenhang mit dem Trendverfahren behandeln. – Für die Lautfolge Sonant–Konsonant stehen uns in Bd. 3 181, in Bd. 6 227 und für die Lautfolge Konsonant–Sonant in Bd. 3 128, in Bd. 6 138 Fälle zur Verfügung.

Die Korrelationen: Sonanten-Melodiewinkel/Konsonanten-Melodiewinkel

In früheren Arbeiten habe ich wiederholt auf die Wichtigkeit der *Homogenität* der Kollektive hingewiesen. Das heißt, man kann ein Material nicht variations- oder korrelationsstatistisch untersuchen, wenn es in sich nicht einheitlich ist. Deshalb wurde es bisher stets in vier wesentlich voneinander verschiedene phonologische Gruppen¹⁰ aufgeteilt. Neben den Korrelationen für alle Sonanten sollen auch diesmal die Korrelationen für betonte und unbetonte Silben (genauer gesagt: für phonologisch betonte und phonologisch unbetonte Silbenträger) getrennt gegeben werden. An sich ist es auch nötig, zwischen Längen und Kürzen zu unterscheiden. Doch können wir hier eine Zusammenfassung vornehmen, da die Differenzen nicht groß sind. Gerade bei Korrelationen ist es auch wichtig, das Material nicht durch allzu starke Aufteilung zu klein zu machen. Deshalb habe ich mich bei der vorliegenden Arbeit nur für die Zweiteilung in «betont» und «unbetont» entschieden.

Es gibt allerdings hinsichtlich der Silbenmelodie einige charakteristische Unterschiede zwischen Längen und Kürzen. Auf diese werden wir sogleich zu sprechen kommen.

⁹ Es gibt allerdings auch «dreidimensionale» Korrelationen (vgl. E. Czuber). Doch erfordern diese einerseits eine sehr umständliche mathematische Berechnung, andererseits sind die Ergebnisse weniger anschaulich als bei der zweidimensionalen Korrelation, die wir anwenden.

^{9a} Stimmhafte Konsonanten im Silbenschnitt (nach kurzen betonten Sonanten) wurden sowohl mit dem vorhergehenden wie auch mit dem folgenden Sonanten als silbenbildend angesehen.

¹⁰ Lang betont, lang unbetont, kurz betont, kurz unbetont.

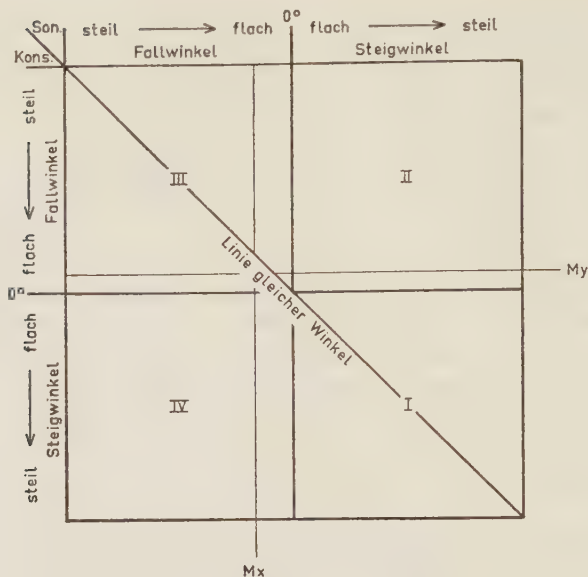


Abb. 1. Schema der Korrelation:
Sonanten-Melodiewinkel / Konsonanten-Melodiewinkel
für die Lautfolgen:
Sonant-Konsonant und Konsonant-Sonant

Erklärung: M_x = arithm. Mittel der Sonanten-Steig- und -Fallwinkel
 M_y = arithm. Mittel der Konsonanten-Steig- und -Fallwinkel
 I und III = positive Quadranten
 II und IV = negative Quadranten, durch das Korrelations-
 kreuz M_x M_y voneinander getrennt

Das Korrelationsschema (Abb. 1) möge die Verteilung der Sonanten- und Konsonantenwinkel auf dem Korrelationsnetz erläutern. Treffen Steigwinkel der Sonanten mit solchen der Konsonanten zusammen, so liegen sie innerhalb des (positiven) Quadranten I. Treffen Fallwinkel der Sonanten mit solchen der Konsonanten zusammen, so liegen sie innerhalb des (positiven) Quadranten III. Je mehr die Gradzahl der Sonanten- mit der der Konsonantenwinkel übereinstimmt, um so näher liegen die Fälle an der «Linie gleicher Winkel». Treffen Steigwinkel der Sonanten mit Fallwinkeln der Konsonanten zusammen, so liegen die Fälle im (negativen) Quadranten II. Treffen Fallwinkeln der Sonanten mit Steigwinkeln der Konsonanten zusammen, so liegen die Fälle im (negativen) Quadranten IV. Positiv nennt man die Korrelation, wenn die Fälle in den Quadranten I und III, negativ, wenn die

Fälle in den Quadranten II und IV überwiegen, wobei noch eine Rolle spielt, ob die Fälle extrem oder nahe dem Mittelpunkt des Korrelationsnetzes angeordnet sind. Im Höchsthalle ist der Koeffizient r der Korrelation $= \pm 1$. Bei gänzlicher Beziehungslosigkeit beider Variablen ist $r = 0$ ¹¹.

Für die verschiedenen Korrelationen wurden die in Tabelle I zusammengestellten Koeffizienten errechnet.

Tab. I

Koeffizienten der Korrelationen zwischen Sonanten- und Konsonantenwinkeln

			Lautfolge: Sonant-Kons.	Lautfolge: Kons.-Sonant
Bd. 3	alle	Silben	+ 0,19	+ 0,14
	bet.	„	+ 0,27	+ 0,09
	unbet.	„	+ 0,09	+ 0,22
Bd. 6	alle	„	+ 0,28	+ 0,19
	bet.	„	+ 0,26	+ 0,05
	unbet.	„	+ 0,10	— 0,15

Wie die Tabelle zeigt, sind die Koeffizienten allgemein durchaus niedrig, teilweise *sehr* klein, in einem Falle sogar negativ. Letzteres würde bedeuten: Steig- und Fallwinkel der Sonanten sind denen der Konsonanten eher entgegengesetzt als gleichlaufend. — Mit einer Ausnahme ist die Abhängigkeit zwischen Sonanten- und Konsonantenwinkeln bei den unbetonten Sonanten noch geringer als bei den betonten. Wir werden darauf weiter unten noch näher zu sprechen kommen. Diese Tatsachen sprechen also allein schon nicht dafür, daß die Silbenmelodie im allgemeinen geradlinig ist.

Im einzelnen muß auffallen, daß der Korrelationskoeffizient für alle Silben zusammen in Bd. 6 höher ist als für betonte und unbetonte, und zwar in beiden Lautfolgen. Normalerweise sollte der r -Wert für *alle* Sonanten dazwischenliegen (wie in Bd. 3). Der Grund dafür liegt in der bereits erwähnten Inhomogenität des Gesamtmaterials. Es ist nämlich sehr wichtig, daß die Kreuzmitten zweier Korrelationen (vgl. Abb. 1), die addiert werden sollen (wie in unserem Falle betonte und unbetonte Silben) annähernd an der gleichen Stelle liegen, m. a. W.: die M_x - und M_y -Linien beider

¹¹ Vgl. die in Anm. 9 erwähnte Schrift von E. Czuber.

Einzelkorrelationen müssen sich möglichst decken. Ist das nicht der Fall, so würde man bei deren Addition etwa in denselben Fehler verfallen, als wenn man beispielsweise zwei völlig gleiche Korrelationen, deren richtige Addition natürlich denselben Wert wie für jede einzelne Korrelation ergibt, unter Verschiebung der Kreuzmitten addieren wollte, wie es unsere schematische Abbildung 2 zeigt, wo bei jeder Einzelkorrelation $r = 0$ ist, während die Addition einen Korrelationskoeffizienten von $+0,27$ ergibt.

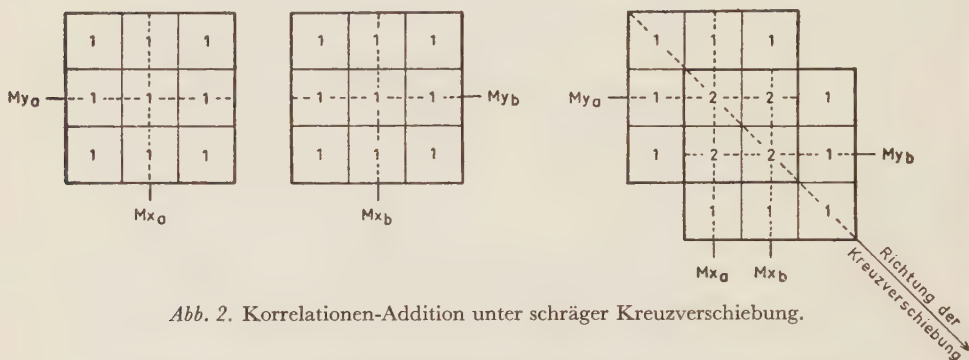


Abb. 2. Korrelationen-Addition unter schräger Kreuzverschiebung.

Ganz allgemein gilt folgendes: Sind die Kreuzmitten in Richtung links oben–rechts unten (wie in der Zeichnung) gegeneinander verschoben, so ist durch die Addition die Tendenz zum Positiven (größerem positiven bzw. kleinerem negativen r) gegeben. Sind die Kreuzmitten in der entgegengesetzten Richtung, also rechts oben–links unten, gegeneinander verschoben, so ist, wie leicht ersichtlich, durch die Addition die Tendenz zum Negativen (größerem negativen bzw. kleinerem positiven r) gegeben. – Werden dagegen bei der Addition die Kreuzmitten waagrecht bzw. senkrecht verschoben, m. a. W.: fallen nur die beiden M_x -Linien bzw. nur die beiden M_y -Linien nicht zusammen, während die jeweils anderen sich decken, so ändert sich an dem Wert der Korrelation nichts: wie aus Abbildung 3 ohne nähere Erläuterung zu ersehen ist. Dies ist für das Verständnis des folgenden wichtig.

Betrachten wir daraufhin die Lage der Korrelationskreuze bei den betonten und bei den unbetonten Silben, so finden wir, daß das *Sonanten*mittel der betonten von dem der unbetonten sehr verschieden ist, und zwar bei beiden Sprechern: natürlich, weil die betonten Sonanten vorwiegend steigend, die unbetonten vorwiegend fallend sind. Hinsichtlich der *Konsonanten* dagegen verhalten sich

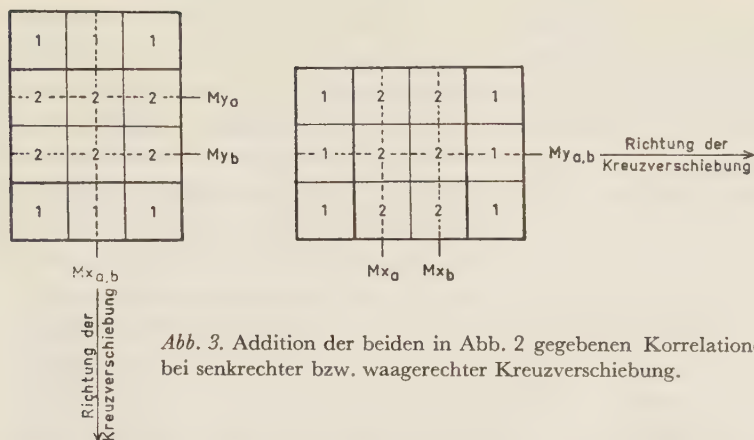


Abb. 3. Addition der beiden in Abb. 2 gegebenen Korrelationen bei senkrechter bzw. waagerechter Kreuzverschiebung.

die beiden Sprecher nicht gleich: In Bd. 3 liegt das Konsonantenmittel in betonten Silben von dem in unbetonten nur sehr wenig entfernt, m. a. W.: die M_y -Linien fallen annähernd zusammen. Gemäß Abbildung 3 ändert sich daher bei der Addition der beiden Einzelkorrelationen nicht viel, und der Korrelationskoeffizient für alle Silben liegt in Bd. 3 deshalb annähernd in der Mitte zwischen dem für die betonten und dem für die unbetonten Silben.

In Bd. 6 dagegen bestehen auch bei den Konsonanten große Verschiedenheiten: In der Lautfolge Sonant-Konsonant sind die Konsonantenmittel bei den betonten Silben gegenüber den unbetonten um mehr als 5° , in der Lautfolge Konsonant-Sonant sogar um mehr als 10° nach der positiven Seite hin verschoben¹². Deshalb ist bei dieser Lautfolge der Korrelationskoeffizient für *alle* Silben in Bd. 6 auch besonders weit von dem für betonte und unbetonte Silben allein entfernt. Auf die *Gründe* für die Verschiedenheit bei den beiden Sprechern einzugehen, ist hier nicht der Platz. – Es braucht nicht erst hervorgehoben zu werden, daß die Koeffizienten der Einzelkorrelationen unter diesen Umständen den wirklichen Verhältnissen mehr gerecht werden als der Koeffizient für alle Silben, womit sich die Notwendigkeit der Trennung in betonte und unbetonte Silben erwiesen hat.

Für die Addition von Längen und Kürzen gelten natürlich dieselben Grundsätze. Gegenüber der Opposition betont-unbetont

¹² Da das Sonantenmittel in den betonten Silben gegenüber dem in unbetonten auch viel mehr nach dem Positiven hin liegt, differiert die Kreuzlage bei betonten und unbetonten Silben also in der in Abb. 2 angedeuteten Richtung, d. h. die Gesamtkorrelation tendiert mehr zum Positiven hin.

bestehen hier wesentliche Unterschiede: Was die Sonanten betrifft, so sind lange im Durchschnitt flacher als kurze¹³. Da dies nicht nur für betonte und unbetonte Sonanten, sondern auch für steigende und fallende Melodiewinkel in annähernd gleichem Maße gilt, so geht schon daraus hervor, daß die M_x -Linien der Sonantenmittel bei langen und kurzen Sonanten nur wenig voneinander differieren können. Da obendrein die Lage bei den Konsonanten nicht viel anders ist, kann man, wie oben bereits erwähnt, Silben mit langen und Silben mit kurzen Sonanten ohne Bedenken zusammenfassen.

Trotzdem gibt es, wie ebenfalls oben angedeutet wurde, einige Unterschiede zwischen Längen und Kürzen. Diese hängen mit der *Form* der Sonoren zusammen. Nach steigenden *kurzen* Sonanten setzt sich nämlich der Anstieg oft noch bis in den folgenden Konsonanten hinein fort. Bei steigenden *langen* Sonanten dagegen ist der melodische Höhepunkt der Silbe meist schon innerhalb des Sonanten erreicht^{13a}. Deshalb hat der Konsonant nach steigenden langen Sonanten überwiegend fallende Tendenz (vgl. Abb. 4, Form III), und zwar in beiden Textlisten gleichmäßig zu 73 %. Nach kurzen steigenden Sonanten hat der Konsonant dagegen im Durchschnitt – bei geringer Abweichung zwischen den beiden Sprechern – noch zu 37 % steigende Gesamttendenz (Abb. 4, Form I), und wo er fällt,



Abb. 4. Schematische Darstellung häufiger Silbenmelodieformen bei kurzen und langen Sonanten.

ist der Winkel doch flacher als nach langen Konsonanten, weil der erste, steigende Teil des Konsonanten auf die Gesamttrichtung seines Melodieverlaufs verflachend wirkt (Abb. 4, Form II). Hieraus geht hervor, daß bei den Kürzen der Quadrant I stärker bzw. der Quadrant II schwächer belegt sein muß als bei den Längen (vgl. Abb. 1), was wiederum zur Folge hat, daß die Korrelation der Kürzen, da der Quadrant I «positiv», Quadrant II «negativ» ist (s. o.), einen höheren positiven Koeffizienten haben wird.

¹³ Vgl. A. Maack, 1955, S. 294 ff.

^{13a} Vgl. A. Maack, 1957, bei Anm. 137–141.

Neben diesem Unterschied zwischen Längen und Kürzen bei steigender Sonantenmelodie gibt es noch einen etwas weniger deutlichen bei fallender Sonantenmelodie. Es gilt nämlich der Satz: Nach fallenden kurzen Sonanten sind steigende Konsonanten häufiger als nach fallenden langen Sonanten. Dieser «Richtungswechsel», wie wir das Umschlagen von steigender zu fallender und von fallender zu steigender Melodie in den Lautfolgen Sonant-Konsonant und Konsonant-Sonant fortan nennen wollen, ist in Bd. 6 etwas seltener als in Bd. 3, aber bei beiden Sprechern ist nach fallenden kurzen Sonanten der Richtungswechsel häufiger als bei langen: im Durchschnitt beider Texte nach kurzen Sonanten 25%, nach langen 15%. (Wie der Vergleich mit den eben gegebenen Zahlen erweist, tritt also der Richtungswechsel nach fallenden Sonanten ganz erheblich seltener auf als nach steigenden.)

Da bei kurzen Sonanten der «negative» Quadrant IV (s. Abb. 1) demnach etwas stärker belegt ist als bei langen, geht der Einfluß auf den r-Wert natürlich dahin, daß bei kurzen Sonanten eine leichte negativierende Tendenz besteht, so daß die durch den ersten (in Abb. 4 erläuterten) Satz gegebene Tendenz wieder abgeschwächt wird. Im ganzen liegen daher zwischen den Korrelationskoeffizienten der Längen und denen der Kürzen weder in Bd. 3 noch in Bd. 6 nennenswerte Differenzen vor.

Hinsichtlich der Lautfolge Konsonant-Sonant waren ähnliche Unterschiede zwischen Längen und Kürzen nicht erkennbar. Entsprechendes gilt demgemäß auch für die Koeffizienten der Einzelkorrelationen dieser Lautfolge.

«Richtungswechsel» und «Winkelnormen»

Mit dem Richtungswechsel haben wir bereits eines der charakteristischsten Gesetze der Silbenmelodie berührt. Der Richtungswechsel ist einer der Hauptgründe dafür, daß die Koeffizienten aller Korrelationen so sehr klein sind. Denn er gilt nicht nur für die soeben etwas näher behandelte Lautfolge Sonant-Konsonant, sondern in fast genau demselben Maße auch für die Folge Konsonant-Sonant.

Die Häufigkeit des Richtungswechsels für die beiden Lautfolgen gibt die Tabelle II in Prozentzahlen¹⁴. Aus der Tabelle geht

¹⁴ Die wenigen Fälle mit «gehaltener» Melodie, d. h. wo die Melodie des Sonanten bzw. Konsonanten, im ganzen gesehen, weder steigt noch fällt, wurden bei der Berechnung nicht berücksichtigt.

Tab. II

Prozentzahlen des Richtungswechsels

Erkl.: Die Ausdrücke «steigend» und «fallend» beziehen sich auf die *Sonantenmelodie*

				Lautfolge: Sonant-Kons.	Lautfolge: Kons.-Sonant	
Bd. 3	{	alle	{	steigend	62	60
		Silben	{	fallend	32	27
		bet.	{	steigend	59	61
		Silben	{	fallend	30	28
		unbet.	{	steigend	77	57
		Silben	{	fallend	34	27
		Gesamtdurchschnitt			45	41
Bd. 6	{	alle	{	steigend	68	40
		Silben	{	fallend	15	35
		bet.	{	steigend	65	30
		Silben	{	fallend	13	63
		unbet.	{	steigend	81	87
		Silben	{	fallend	15	27
		Gesamtdurchschnitt			39	37


hervor, daß in nicht viel weniger als in der Hälfte aller Fälle der Konsonant eine dem vorhergehenden bzw. folgenden Sonanten entgegengesetzte Gesamtmelodierichtung hat, besonders in Bd. 3. In Bd. 6 ist der Durchschnitt um etwa 5% geringer. Ob individuelle Unterschiede dafür maßgebend sind, ob es an dem Thema liegt (Erzählung-Vorlesetext), oder ob es sich nur um einen Zufall handelt (die geringe Differenz von 5% schließt diese Möglichkeit nicht aus), kann hier nicht entschieden werden. In der Lautfolge Konsonant-Sonant ist der Richtungswechsel bei beiden Sprechern um eine Kleinigkeit seltener (in Bd. 6 bei den Fallwinkeln dagegen häufiger).


Wie schon soeben erwähnt wurde, ist der Richtungswechsel bei steigenden Sonanten weitaus öfter als bei fallenden: wie wir aus der Tabelle erschen, sowohl in betonten wie auch in unbetonten Silben. Dabei übersteigt der Prozentsatz der unbetonten Silben sogar noch den der betonten: hauptsächlich bei den steigenden, in geringem Maße aber auch bei den fallenden Sonanten. Da der Richtungswechsel, wie wir gesehen haben, negativierend auf den Korrelationskoeffizienten wirkt, haben wir nunmehr auch eine Erklärung


dafür, daß die r-Werte der unbetonten Silben noch weniger positiv sind als die der betonten.


Das alles gilt ohne Einschränkung jedoch nur für die Lautfolge Sonant–Konsonant bei beiden Sprechern. In der Lautfolge Konsonant–Sonant sind die Gesetze anscheinend weniger fest, wie ja auch der Prozentsatz des Richtungswechsels etwas geringer ist. In Bd. 3 sind die Verhältnisse zwischen betonten und unbetonten Silben, wenigstens bei den Sonanten mit steigender Melodie, sogar umgekehrt. Bezeichnenderweise ist deshalb auch der Korrelationskoeffizient der unbetonten Silben in Bd. 3 mehr positiv als der der betonten: was wir oben als einzige Ausnahme von der Regel kennengelernt hatten. – In Bd. 6 ist der Richtungswechsel in den unbetonten Silben zwar im Durchschnitt größer als in den betonten, namentlich bei steigenden Sonanten, weshalb auch der Korrelationskoeffizient der unbetonten Silben in Bd. 6 ganz besonders zum Negativen tendiert (vgl. Tab. I). Dafür finden wir aber bei den betonten Silben die einzige – und zwar sehr deutliche – Ausnahme von der Regel, daß der Richtungswechsel bei den fallenden Sonanten seltener ist als bei den steigenden. Um einen Zufall kann es sich kaum handeln. Die betonten Silben mit fallender Sonantenmelodie sind zwar mit 19 Fällen ziemlich selten, die steigenden sind aber mit 40 vertreten, und deren Prozentsatz (30) ist, verglichen mit dem in allen anderen Gruppen der Lautfolge, außerordentlich niedrig. Eine Erklärung für diese Erscheinung kann hier nicht gegeben werden. Sie stimmt aber überein mit dem, was wir im folgenden Abschnitt über diese Gruppe zu sagen haben.

Der Richtungswechsel ist zwar, besonders bei steigenden Sonanten, eine recht häufige Erscheinung. Immerhin gibt es doch aber noch ziemlich viele Ausnahmen. Im Grunde genommen ist der Richtungswechsel nur die extreme Form einer Erscheinung, die mit noch viel geringeren Ausnahmen gilt – auch für *fallende* Sonantenmelodie – und die ich deshalb als die «*Winkelnorm*» der Silbenmelodie bezeichnen möchte. Haben wir in der Lautfolge Sonant–Konsonant bei steigenden Sonanten häufig diese Form (als


Richtungswechsel):  S K, so ist die «abgemilderte» Form diese:



 S K, d. h. der Konsonant braucht nicht gerade zu fallen, aber er steigt weniger. Ähnlich wäre für die extreme Form bei steigenden

Sonanten der Lautfolge Konsonant–Sonant:  die abgemil-

derte Form diese:  **K S**, d. h. auch in diesem Falle steigt der Konsonant weniger als der Sonant. – Entsprechendes gilt für die Fallwinkel. Die (verhältnismäßig seltene) Extremform der Folge

Sonant–Konsonant ist diese:  **S K**. Die abgemilderte Form:

 **S K** besagt, daß der Konsonant nach fallendem Sonanten weniger fällt als dieser. Und ähnlich bei der Lautfolge Konsonant–Sonant:

Für die Extremform:  **K S** ist die abgemilderte Form diese:  **K S**; d. h. auch hier fällt der Konsonant weniger als der Sonant. Die soeben schematisch angedeuteten Formen können wir unter Ein- schluß des Richtungswechsels als Winkelnormen der Silbenmelodie in folgende Regeln zusammenfassen: *Steigt die Sonantenmelodie, so ist die Melodie des – vorhergehenden oder folgenden – Konsonanten meist weniger steigend oder sogar fallend. Fällt die Sonantenmelodie, so ist die Melodie des – vorhergehenden oder folgenden – Konsonanten meist weniger fallend oder sogar steigend.* Fälle, wo der Konsonant stärker steigt oder stärker fällt als der vorhergehende oder folgende Sonant, sind deshalb als *Anomalitäten* zu bezeichnen. – War bei den fallenden Sonanten der Richtungswechsel oft noch eine Ausnahme, so gelten die Winkelnormen auch für diese mit verhältnismäßig geringen Einschränkungen. Wie die Tabelle III zeigt, machen die Anomalitäten selbst im ungünstigsten Falle nur etwa ein Drittel aus^{14a}.

Im übrigen gilt in ähnlicher Weise das, was über den Richtungswechsel festgestellt werden konnte: In der Lautfolge Sonant–Konsonant ist die Regel etwas fester als in der Folge Konsonant–Sonant: in Bd. 6 bei den Fallwinkeln dagegen wieder weniger fest. Bei steigenden Sonanten gilt die Regel strenger als bei fallenden: wieder mit der einen Ausnahme der betonten Silben in der Lautfolge Konsonant–Sonant in Bd. 6. Bei den unbetonten Silben gibt

^{14a} Die verhältnismäßig wenigen Fälle, wo Sonant und Konsonant den gleichen Steig- bzw. Fallwinkel haben, sind ausgelassen worden.

Tab. III
 Prozentzahlen der Winkelnormen
 Erkl. s. bei Tab. II

				Laufolge: Sonant-Kons.	Laufolge: Kons.-Sonant
Bd. 3	alle	{ steigend		92	89
	Silben	{ fallend		82	65
	bet.	{ steigend		91	90
	Silben	{ fallend		69	56
	unbet.	{ steigend		100	83
	Silben	{ fallend		89	69
	Gesamtdurchschnitt			86	76
Bd. 6	alle	{ steigend		96	71
	Silben	{ fallend		68	87
	bet.	{ steigend		96	66
	Silben	{ fallend		65	95
	unbet.	{ steigend		95	94
	Silben	{ fallend		69	84
	Gesamtdurchschnitt			81	81

es, entsprechend dem Richtungswechsel, im Durchschnitt noch weniger Anomalitäten als bei den betonten: wieder mit Ausnahme der Laufolge Konsonant-Sonant – wenigstens bei den steigenden Sonanten – in Bd. 3. Im ganzen gesehen ist also die Tabelle der Winkelnormen, wie nicht anders zu erwarten war, ein getreues Spiegelbild der Tabelle des Richtungswechsels¹⁵.

Die Zahl der Anomalitäten verringert sich teilweise noch weiter, wenn wir die Fälle an Sprechabschnittsenden (ohne «weiterweisenden Sprechтакт»¹⁶, wo die Stimme eher gehoben als gesenkt wird) fortlassen. Von vornherein ist nämlich anzunehmen, daß an solchen Stellen die Melodie eines Endkonsonanten die Tendenz zu

¹⁵ *Bethge* unterscheidet in seiner oben erwähnten Arbeit (s. Anm. 2) zwischen «konvexen» (/ bzw. \) und «konkaven» (/ bzw. \) Melodieformen der Laufolge Sonant-Konsonant, ohne dabei eine Trennung zwischen Steig- und Fallwinkeln vorzunehmen. Unseren Normen entsprechen demnach die konvexen Steig- und die konkaven Fallwinkel, während die andern beiden Formen nach unserer Definition anomal sind. Die verhältnismäßig großen Unterschiede, die B. hinsichtlich konvexer und konkaver Form zwischen Bd. 5 und Bd. 6 findet (in Bd. 6 37 %, in Bd. 5 nur 20 % konkav), sind hauptsächlich durch den verschiedenen prozentualen Anteil der fallenden und steigenden Winkel, in zweiter Linie wohl besonders durch die verschiedene Prozentzahl der anomalen Fälle bei den Fallwinkeln bedingt, die in den beiden von uns verwendeten Texten auch verschieden ist. Wir erhalten für Bd. 6, wie B., 37 %, für Bd. 3, den B. nicht bearbeitet hat, sogar 50 %.

¹⁶ Vgl. *H. Klinghardt*.

starkem Fallen haben wird, so daß unter Umständen der Winkel steiler als der des vorhergehenden Sonanten werden wird, also anomal. Die Untersuchung ergab für die fallenden Winkel der Lautfolge Sonant–Konsonant an Sprechabschnittsenden in Bd. 3 tatsächlich 53% (gegen 18% für alle Fälle). In Bd. 6 beträgt die Differenz freilich nur wenige Prozent. Hierbei ist aber zu berücksichtigen, daß die Prozentzahl in dieser Textliste ohnehin schon merklich höher ist als in der anderen, und daß ferner Sprechabschnittsenden (ohne weiterweisenden Sprechтакт) in einem Vorlesetext gewöhnlich seltener sind als in einer Erzählung (in unserem Falle nur halb so häufig), so daß die Zahl für statistische Untersuchungen etwas zu knapp wird, um gültige Aussagen zu machen. – Für die Lautfolge Konsonant–Sonant an Sprechabschnittsenden betragen die Unterschiede gegen alle Fälle in beiden Textlisten nur ein paar Prozent. Hier ist ja ein Einfluß der Stellung im Satz auch kaum zu erwarten.

Im umgekehrten Sinne scheint die Frage und der weiterweisende Sprechтакт zu wirken, soweit an solchen Stellen die Stimme merklich gehoben wird: Folgt auf einen Silbenträger dann noch ein stimmhafter Konsonant, so scheint dessen Melodie verhältnismäßig oft noch stärker zu steigen als die Sonantenmelodie. Genauere Zahlen können hierfür jedoch nicht gegeben werden, da zu wenig solcher Fälle vorliegen.

Die dreisonorige Silbe

Die vorstehenden Ausführungen gelten, wie eingangs erwähnt, sowohl für zweisonorige wie für dreisonorige Silben. Es bleibt nur noch übrig, die dreisonorige Silbe als *Einheit* zu behandeln. Das Material ist natürlich erheblich kleiner. Es stehen uns aber im ganzen doch etwa 50 Fälle in beiden Texten zur Verfügung, die als Ergänzung zu unseren bisherigen Darlegungen ausreichen dürften.

Aus dem, was in dem vorigen Kapitel über Winkelnormen und Richtungswechsel für die beiden Lautfolgen gesagt wurde, kann man auf die Form der dreisonorigen Silbe bereits Schlüsse ziehen.

Die Abbildung 5 gibt eine schematische Übersicht über die vorkommenden Melodieformen der Lautfolge Konsonant–Sonant–Konsonant. Die Konsonantenmelodie ist teilweise gestrichelt wiedergegeben. Es sollte damit offengelassen werden, ob die «abgemilderte» oder die «extreme» Form, d.h. Richtungswechsel, sei es bei

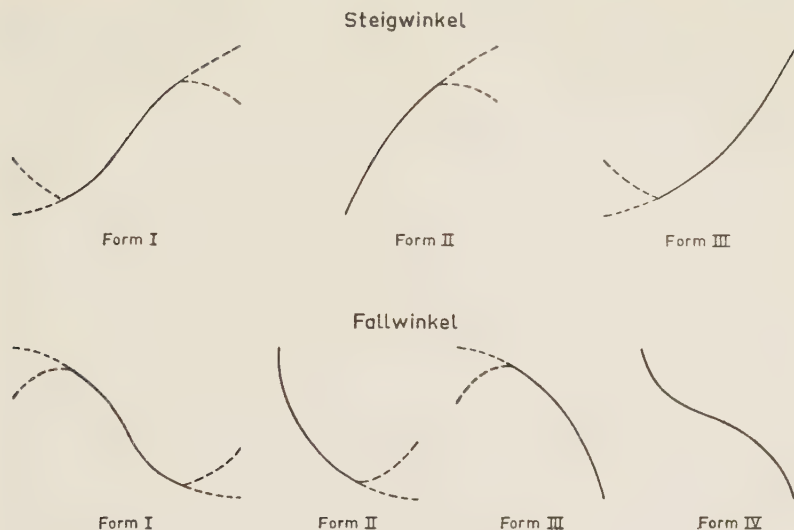



Abb. 5. Melodieformen der dreisonorigen Silben.

der Lautfolge Sonant–Konsonant, sei es bei der Folge Konsonant–Sonant, vorliege.

Wie zu erwarten, ist bei den Steigwinkeln die Form I, die den Winkelnormen entspricht, in beiden Textlisten die häufigste. Die selten vorkommende Form II der Steigwinkel entspricht den Anomalitäten in der Lautfolge Konsonant–Sonant: Hier ist der Winkel des *Anfangs*konsonanten steiler als der des Sonanten. Die auch nicht häufige Form III der Steigwinkel entspricht den Anomalitäten in der Lautfolge Sonant–Konsonant, mit steilerem *End*konsonanten. Außerdem gibt es noch einige Fälle mit *mehrfacher* Biegung. Die

doppelt anomale Form:  ist bezeichnenderweise weder in Bd. 3 noch in Bd. 6 vorhanden.

Bei den Fallwinkeln ist die den Normen entsprechende Form I wieder die Hauptform. Sie kommt aber wegen der bei den Fallwinkeln etwas häufigeren Anomalitäten in beiden Textlisten etwas seltener vor als die Form I bei den Steigwinkeln. Die Form II entspricht den Anomalitäten in der Folge Konsonant–Sonant, die Form III den Anomalitäten in der Folge Sonant–Konsonant. Beide sind erheblich seltener als die Hauptform. Form IV entspricht der doppelten Anomalität. Sie kommt in Bd. 3 mehrmals vor, in Bd. 6

nicht. Daneben gibt es noch einige Fälle mit mehrfacher Biegung, wie bei den Steigwinkeln.

Über die Häufigkeit der einzelnen Formen in den beiden Textlisten für sich und im Durchschnitt beider Textlisten soll die Tabelle IV unterrichten. Wie aus der Tabelle zu ersehen, kommen manche Nebenformen nur in einer der beiden Textlisten vor. Bei dem etwas knappen Material kann es natürlich nicht überraschen, wenn die Abweichungen zwischen den beiden Sprechern diesmal – scheinbar – etwas größer sind. Die *Grundtendenz* ist aber bei beiden offenbar dieselbe. – Dazu ist noch zu bemerken, daß die charakteristischen Formen:

 bzw.  oft auch in den Fällen gewahrt

sind, die in der Tabelle IV zu den Nebenformen gezählt wurden, d. h., wo Anomalitäten vorliegen. Zum Beispiel kommt es vor, daß die Melodie des ersten Konsonanten im ganzen stärker steigt als die des folgenden Sonanten, aber trotzdem mit flachem Einsatz beginnt. Entsprechendes gilt für die Fallwinkel.

In meinen beiden ersten Melodiarbeiten (vgl. Anm. 1) hatte ich mich mit der Darstellung der Lautmelodiekurven nach dem Trendverfahren beschäftigt. Der Trend¹⁷ ist eine – gerade oder gekrümmte – Linie, die sich einem gegebenen Kurvenzuge möglichst genau anpaßt, d. h. die mittlere (quadratische) Abweichung der einzelnen Punkte des Kurvenzuges von der Linie muß ein Minimum werden. Diese Abweichung (im folgenden m_2 genannt) ist also ein Maß für die Güte der Annäherung der Trendlinie an den Kurvenzug: Je größer die Annäherung, um so kleiner ist m_2 . Der Trend 1. Grades ist eine gerade Linie (wie sie zur Bezeichnung des Winkels benutzt wurde, den der geradlinig gedachte Melodieverlauf

Tab. IV

Prozentuale Häufigkeit der typischen Melodieformen bei dreisonorigen Silben

	Steigwinkel				Fallwinkel		
	Bd. 3	Bd. 6	D'schn.		Bd. 3	Bd. 6	D'schn.
Form I	53	86	64	Form I	42	55	48
Form II	—	14	5	Form II	—	18	9
Form III	20	—	14	Form III	8	18	13
mehrf. Biegung	27	—	18	Form IV	25	—	13
				mehrf. Biegung	25	9	17

¹⁷ Vgl. P. Lorenz.

mit der – waagerechten – Zeitachse bildet), der Trend 2. Grades eine einfach gekrümmte, der Trend 3. Grades eine doppelt gekrümmte Kurve usw. –

Aus der Form der dreisonorigen Silbe müßte man schließen, daß der Trend 3. Grades eine sehr gute Angleichung bringt. Die mittlere Abweichung des Trends 3. Grades ($m_2^{(3)}$ für Bd. 3 und Bd. 6 ist 0,535 bzw. 0,599. Diese Zahlen sagen nun an sich noch gar nichts, wenn man nicht durch Vergleiche einen Maßstab gewinnt. Diese Vergleiche können auf verschiedene Weise gewonnen werden: Erstens kann man die $m_2^{(3)}$ -Werte (für den Trend 3. Grades) mit den m_2 -Werten der Trends niedrigerer Grade in Beziehung setzen, m. a. W.: man kann die jeweilige «Besserung» berechnen, die ein Trend gegenüber dem Trend des nächstniedrigeren Grades bringt. Es ist auf diese Weise möglich, ziemlich genaue Aussagen über die vorherrschenden Kurvenformen zu machen. Die «Besse-

rung» ist in der vorliegenden Arbeit nach der Formel $\frac{m_2^{(x)}}{m_2^{(x-1)}} \cdot 100$

berechnet worden. Nach dieser Definition ist die «Besserung» also das prozentuale Verhältnis der mittleren Abweichung eines Trends zu der mittleren Abweichung des Trends nächstniedrigeren Grades: je niedriger die Prozentzahl, um so größer daher die «Besserung»¹⁸.

Die Besserungswerte sind

in Bd. 3 f. d. Trend 2. Grades gegenüber d. Trend 1. Grades: 74%,

„ „ 3 „ „ „ 3. „ „ „ „ 2. „ : 77%,

„ „ 6 „ „ „ 2. „ „ „ „ 1. „ : 74%,

„ „ 3 „ „ „ 3. „ „ „ „ 2. „ : 78%.

Die Zahlen entsprechen vielleicht nicht ganz den Erwartungen, weil hiernach doch der Trend 2. Grades, d. h. die *einfach* gekrümmte Kurve, eine noch bessere Angleichung gegenüber dem nächstniedrigeren Trend bringt als die doppelt gekrümmte Kurve, die wir als charakteristisch für die Silbenmelodie herausgestellt hatten. Wir müssen hierbei aber bedenken, daß die beiden Hauptformen

¹⁸ In den früheren Arbeiten wurde die Besserung berechnet nach der Formel $\frac{m_2^{(x)} - m_2^{(x+1)}}{m_2^{(0)}} \cdot 100$. Je größer die Besserung, um so *höher* waren also die Prozentwerte.

Die jetzige Formel, bei der es umgekehrt ist, kommt i. g. zu ziemlich ähnlichen Ergebnissen, hat aber den Vorteil, daß sie kürzer ist und von dem $m_2^{(0)}$ -Wert unabhängig macht.

für Steig- und Fallwinkel auch bei der größten Besserung durch den Trend 2. Grades gewahrt bleiben können: Die Prozentzahlen drücken dann eben aus, daß die Anfangs- bzw. Endkrümmung oft nur verhältnismäßig schwach ist, so daß schon durch eine einfach gekrümmte Kurve eine bedeutende Besserung der Angleichung erfolgt und die Trends höheren Grades keine große Besserung der Angleichung mehr bringen können.

Eine andere Vergleichsmöglichkeit bieten die Trends der Einzellaute¹⁹, die allerdings nur für Bd. 6 vorliegen. Es zeigt sich, daß die $m_2^{(1)}$ -Werte der Einzellaute viel niedriger sind als die der Silben: 0,778 gegen 1,030. Das bedeutet, daß die Melodie der Einzellaute erheblich geradliniger verläuft als die der Silben. Ferner bringt bei den Einzellauten der Trend 2. Grades eine *erheblich* größere Besserung als der Trend 3. Grades (64,8 gegen 84,3%). Dies bestätigt also das, was schon in den ersten Arbeiten festgestellt wurde: daß bei den Einzellauten in den weitaus meisten Fällen ein gerader oder höchstens einfach gekrümmter Melodieverlauf vorliegt, während die Silbenmelodie erst durch den Trend 2. oder 3. Grades im wesentlichen erfaßt wird.

Zum Schluß noch ein Wort über den Einfluß der *Quantität* auf die Form der Silbenmelodie. Hierzu ist im ganzen folgendes zu sagen: Mit zunehmender Silbendauer wird die Abweichung größer: je höher der Trendgrad, um so mehr. Die Koeffizienten der Korrelation m_2/q ²⁰ sind nämlich

für den Trend	1. Grades	in Bd. 3	= +0,33,	in Bd. 6	= +0,29
„ „ „	2. „ „ „	3	= +0,42,	„ „ 6	= +0,40
„ „ „	3. „ „ „	3	= +0,47,	„ „ 6	= +0,49

Mit zunehmender Dauer wird also die Angleichung des Melodiekurvenzuges an die Trendlinien immer schlechter. Das bedeutet, daß die längeren Silben größere Melodieschwankungen, also mehrfache Biegungen, aufweisen. Gerade bei diesen mehrfach

¹⁹ Zum Vergleich der Silbenmelodie mit der Lautmelodie sind von den Einzellauten die kürzesten fortzulassen, da sich für diese die natürlichen Vergrößerungen der Kurve (Festlegung auf $1/100$ sec-Wert) viel stärker bemerkbar machen müssen als für die langen. Man muß bedenken, daß z. B. bei 4 gegebenen Punkten schon die *völlige* Angleichung durch den Trend 3. Grades in jedem Falle erfolgt, da 4 Punkte stets durch eine doppelt gekrümmte Kurve verbunden werden können. Deshalb sind für den Vergleich alle kurzen Einzellaute bis zur Dauer von $9/100$ sec fortgelassen worden: Bei dieser Dauer liegen die kürzesten Silben in Bd. 3 und Bd. 6.

²⁰ q = Silbendauer in $1/100$ sec.

gebogenen Melodiekurven spielt es keine große Rolle mehr, ob die Angleichung durch den Trend 1., 2. oder 3. Grades versucht wird: die Abweichung ist bei letzterem nicht mehr viel geringer als bei dem Trend 2. oder sogar 1. Grades. Da der Unterschied zwischen den einzelnen Trendgraden bei kurzen Silben aber sehr erheblich ist, muß die Korrelation m_2/q gerade bei den höheren Trendgraden einen besonders hohen positiven Koeffizienten aufweisen.

Grundsätzlich gilt dasselbe natürlich auch für die Einzellaute, die wir noch einmal zum Vergleich heranziehen wollen²¹. Trotzdem sind die Verhältnisse hier wesentlich andere. Die Koeffizienten derselben Korrelation sind nämlich

für den Trend	1. Grades	=	— 0,18
„ „ „	2. „	=	— 0,05
„ „ „	3. „	=	+ 0,04.

Praktisch ist also bei den Trends 2. und 3. Grades gar kein Einfluß der Länge festzustellen. Daß der Korrelationskoeffizient für den 1. Grad leicht negativ ist, m. a. W.: daß die Angleichung an die Gerade mit zunehmender Länge besser wird, kommt daher, daß die längeren Einzellaute – im Gegensatz zu den längeren Silben – einen mehr geraden Melodieverlauf haben. Dies gilt allerdings nicht für alle Laute. Gerade bei den Einzellauten sind die Unterschiede in dieser Beziehung recht erheblich, worauf ich schon früher hingewiesen habe²².

Soviel über grundsätzliche Fragen der Silbenmelodie. Die hier herausgestellten Regeln gelten natürlich zunächst nur für das Deutsche, wobei es auch noch nicht ganz sicher ist, ob die deutschen Mundarten sich einbeziehen lassen. Desgleichen muß einer späteren Untersuchung vorbehalten bleiben, ob die oben behandelten Regeln auch für die Melodie eines *größeren* Gefüges Gültigkeit haben.

Zusammenfassung

Anhand von zwei deutschen Schallplattentexten wird das Melodieverhältnis der stimmhaften Laute einer Silbe zueinander (d. h. des Sonanten als Silbenträger und des *unmittelbar* vorhergehenden oder folgenden stimmhaften Konsonanten) untersucht. Zwischen

²¹ Auch hier erst oberhalb $\frac{9}{100}$ sec, aus den gleichen Gründen wie oben.

²² Vgl. besonders Arch. vergl. Phon. 2, S. 152 f., 1938.

Sonanten- und Konsonantenmelodie bestehen sehr große Unterschiede sowohl hinsichtlich der Steilheit als auch hinsichtlich der Richtung (steigend-fallend). Die Melodie des Konsonanten hat oft die entgegengesetzte Richtung wie die des Sonanten, bei steigender Sonantenmelodie sogar überwiegend. Haben Sonant und Konsonant die gleiche Melodierichtung, so ist die Melodie des Konsonanten meist flacher als die des Sonanten. Es macht dabei im allgemeinen keinen sehr großen Unterschied, ob der Konsonant vor oder hinter dem Sonanten steht. Silben mit drei stimmhaften Lauten (mit dem Sonanten in der Mitte) haben demnach in der Regel eine doppelt gekrümmte Melodieform: bei steigender Sonantenmelodie S-förmig, bei fallender umgekehrt S-förmig. Im Gegensatz dazu ist die *Lautmelodie* überwiegend geradlinig oder höchstens *einfach* gekrümmt.

Summary

An investigation was made, using two German gramophone discs as material, into the mutual frequency relations of the voiced sounds of a syllable (i.e. the relations between the syllabic sonant and the immediately preceding or following voiced consonants). Considerable differences were found between the frequency patterns of sonants and those of consonants, both in steepness and in direction (rising-falling). The consonant pattern often has the opposite direction from the sonant, and overwhelmingly so when the sonant pattern is rising. If both sonant and consonant have the same frequency direction, then the consonantal pattern is usually flatter than that of the sonant, and in this respect it makes very little difference whether the consonant precedes or follows the sonant. As a rule, therefore, syllables with three voiced sounds (with the sonant in the middle) have a double change in their frequency pattern, giving an S-like shape for patterns with rising sonant and a reversed S when the sonant is falling. The patterns of *sounds*, on the other hand, are predominantly uni-directional or have at the most a *single* change of direction.

Résumé

Sur la base de deux textes allemands enregistrés, on examine les rapports réciproques des sons voisés d'une syllabe donnée (c'est-à-dire de la sonante en tant que centre de syllabe et de la consonne sonore qui précède ou qui suit immédiatement). Il y a de très grandes différences entre la mélodie de la sonante et celle de la consonne, en ce qui concerne la pente aussi bien que la direction (ascendante-descendante). La mélodie de la consonne a souvent une direction opposée à celle de la sonante et c'est ce qui arrive généralement quand la sonante a une mélodie ascendante. Si la direction est la même pour la consonne et la sonante, la mélodie de la consonne est plus plate que celle de la sonante. Dans l'ensemble, le fait que la consonne précède ou suit la sonante n'a pas grande importance. Les syllabes formées de trois sons voisés (avec la sonante au milieu) ont par conséquent une forme mélodique à double courbure: en forme de S quand la mélodie de la sonante est ascendante, en forme de S renversé quand elle est descendante. Par contre, la mélodie du son isolé est de façon générale rectiligne ou tout au plus à courbure unique.

Literaturverzeichnis

- Bethge, W.: Das Abhören von Lautmelodie und Silbenmelodie. Z. Phon. 7: 320, 1953.
 Czuber, E.: Die statistischen Forschungsmethoden. Wien 1928.
 Klinghardt, H.: Sprechmelodie und Sprechtakt. Marburg 1925,
 Lorenz, P.: Der Trend. Vjh. Konjunkt. Forsch., Sonderh. 9, Berlin 1928.
 Maack, A.: Zum Melodieverlauf nhd. Laute. Arch. vergl. Phon. 2, 145, 1938.
 Id.: Formen des Melodieverlaufs nhd. Laute. Arch. vergl. Phon. 3, 27, 1939.
 Id.: Quantität und Melodiebewegung. Z. Phon. 8, 294, 1955.
 Id.: Die phonometrischen Normen und Methoden und ihre Stellung zur Phonologie.
 Phonetica 1, 82, 184, und 241, 1957.
 Zwirner, E. und K. Zwirner: Phonometrische Forschungen, Reihe B, Bd. 1, 1936.

Adresse des Autors: A. Maack, Schunterstraße 8, Braunschweig (Deutschland)

Froeschels, E.: Phonetica 2: 219–225 (1958).

Zur Frage der Gleichheit von Kau- und Sprechbewegungen

Von EMIL FROESCHELS, New York

Der Kaumethode (Chewing Approach) liegt die Idee zugrunde, daß die Kau- und die Artikulationsbewegungen gleich sein müssen, weil sie einander nicht hindern, wenn man während des Kauens spricht. Denn man kann nicht gleichzeitig zwei verschiedene Bewegungen mit einer Muskelgruppe ausführen. In meinen verschiedenen Veröffentlichungen habe ich absichtlich die Literatur vernachlässigt, weil in der Literatur wohl einigemal die Vermutung geäußert wurde, daß die Artikulationsbewegungen den Kaubewegungen verwandt sind oder von ihnen aus allmählich gebildet wurden, während ich radikal die Identität beider behauptete. Eine Reihe von Veröffentlichungen über diese Ansicht und die Verwendung des lauten Kauens zur Heilung von Stimmstörungen, besonders der funktionellen und verschiedener Sprachstörungen, liegt vor^{1–10}. Sie befassen sich hauptsächlich mit den guten Erfolgen der Kaumethode, abgesehen von Weiß' Beitrag zu dem Buche *The Chewing Approach in Speech Therapy*, der «Chewing and the Origin of Speech» betitelt ist. Auf diese sechzehn Abhandlungen soll erst später zurückgekommen werden. Vorerst sollen die Einwendungen gegen die Kautheorie, insofern sie die Einheit von Kau- und Sprech-

bewegungen betrifft, besprochen werden. Sie erfolgten von drei hervorragenden Fachleuten, der zeitlichen Reihenfolge nach von *Stein*¹¹, *Orthmann*^{12, 13} und *Trojan*¹⁴. *Stein* schreibt, daß die Tätigkeit der Artikulation weder Atmen noch Kauen ist; man könne nur sagen, daß sich diese Tätigkeiten zu Sprechen entwickelt haben und daß die hochentwickelte Funktion des Sprechens deshalb gewisse Eigenheiten des Kauens mit sich trage. *Orthmanns* Ansicht ist, daß «von einer Gleichheit der Funktionen nicht gesprochen werden kann, ...Während kauendes Sprechen den Zahnreihenverschluß meidet und dadurch einen unbefriedigenden Kaueffekt ergibt, betont das sprechende Kauen die Okklusion und behindert den Sprechvorgang beträchtlich. Die Aufmerksamkeit schwankt fortwährend zwischen den beiden Funktionen und vermindert im Effekt beide». Darauf möchte ich sofort einwenden, daß ich ein derartiges Schwanken an mir nicht beobachten konnte, außer wenn ich meine Aufmerksamkeit willkürlich auf Kauen oder auf Sprechen lenkte, wodurch die von mir behauptete Einheitlichkeit eben gestört wurde. Gleichlautende Angaben erhielt ich von sehr zahlreichen Personen, die während des Kauens sprachen. Aber selbst wenn eine Behinderung fühlbar ist, so glaube ich, daß sie durch den Unterschied begründet ist, ob etwas Festes, wie Nahrung, oder Luft im Munde ist. *Trojan*¹⁵, der sich ebenfalls gegen die Einheit der Kau- und Sprechbewegungen äußert, betont, daß «das Sprechen von der Notwendigkeit des Kau-druckes und der Okklusion des Gebisses entlastet scheint». Anderseits sagt er: «Während nach *Wilds* subtiler Beschreibung die Zunge beim Kauen den Bissen den Zähnen auflädt, ihn immer wieder neu formt und zusammenballt, das Vestibulum ausräumt und ihn vor und zurück und von einer Seite auf die andere transportiert, ehe es nach genügender Einspeichelung von der Mitte des Gaumengewölbes zum Schluckakte kommt, sind die Bewegungen beim Sprechen weitgehend symmetrisch.» *Trojan* hat sich in sehr dankenswerter Weise der Mühe unterzogen, die beim lauten Kauen auftretenden Laute auf Band aufzunehmen. Er hat ferner unter Heranziehung linguistischer Probleme und Theorien versucht, die Kau-theorie «etwas zu modifizieren und durch eine Reihe weiterer Argumente zu stützen, woraus sich u. a. auch ein Gewinn für die allgemeine Theorie der Silbe ergibt». Zugleich wird eine Synthese der Lehre von *Froeschels* und der *R. Jakobsons* vom Schichtenbau des Sprachlautsystems angestrebt. In derselben Abhandlung weist *Trojan* darauf hin, daß sich «im entwicklungsgeschichtlichen Rück-

gang die Neigung zu rhythmischem Sprechen verstärkt und sich in dieser Hinsicht der Abstand zwischen Kauen und Sprechen verkürzt». Mit andern Worten, *Trojan*, *Orthmann* und *Stein* zeigen ihr Interesse für die Kautheorie, und keiner von ihnen bezweifelt den Wert der Kaumethode. Ich muß sogar zugeben, daß *Orthmann* und *Trojan* sowie die 16 anderen Autoren mehr für die Popularisierung der Methode und deren Grundlage getan haben als ich selbst. Was speziell *Trojan* und *Orthmann* betrifft, so hat der erstere mit seiner Bandaufnahme und seinen linguistisch-phonetischen Erläuterungen, der letztere mit seinem Buche und den darin enthaltenen Krankengeschichten der Kauidee und Kaumethode einen Grad von Beachtung geschenkt, wie er in der medizinischen Literatur irgendeinem Heilwege nicht häufig zuteil wird. Schon deswegen, aber natürlich auch aus wissenschaftlicher Verpflichtung, möchte ich versuchen, meinen Standpunkt neuerdings und mit zum Teil neuen Beobachtungen und Schlußfolgerungen darzulegen. *Trojan* erwähnt die lateralen Klicks in Negersprachen, die seltenen lateralen L-Laute (z. B. das keltische L) und auch die lateralen Sigmatismen. Was die lateralen Klicks betrifft, so verdanke ich Herrn Dr. *Ernst Westphal*, Lektor für Buschmann- und Hottentottensprachen an der Universität London, briefliche Auskünfte, von denen ich hier nur die die lateralen Klicks betreffende zitiere: «Betreffs Prozentsatz der Schnalze zu den der anderen Konsonanten: Es kommt darauf an, was man zählt und wie man es zählt. Im Wörterbuch könnte man wohl zwischen 20–60% zählen, aber es ist ganz verschieden für verschiedene Sprachen. Ich glaube, daß die vierte Sprache, von mir in meinem „Supplement“ in Vol. III (African Handbook, International African Institute) vermutet, weniger Schnalze hat als Buschmann und Hottentott, aber mehr als Bantu.» Jedenfalls sind laterale Klicks keineswegs Ausnahmereischeinungen. (Sie sind nicht so grundverschieden von den in anderen Sprachen gebräuchlichen Konsonanten, wie es scheinen möchte, denn *Panconcelli-Calcia** hat nachgewiesen, daß sie leicht mit einem folgenden Vokal verbunden werden können.)

Begreiflicherweise bin ich selbst durch Sprachstörungen in meinem Urteil über Lateralisation beeinflusst. Nun ist das seitliche Lispeln eine sehr häufige Erscheinung, so daß ich Tausende der-

* In *Gutzmann, H.*: Physiologie der Stimme und Sprache; 2. Aufl., Seite 200 (Braunschweig 1928).

artiger Fälle sehen konnte. Aber auch laterales L ist nicht selten, kann jedoch oft nur durch genaue optische Beobachtung festgestellt werden, da es sich akustisch kaum vom medialen unterscheidet, außer wenn es stimmlos gesprochen wird. Ferner sah ich einige hörstumme Kinder, die ihre Zunge oft im Munde «herumwälzten». Als sie während der Behandlung zu sprechen anfangen, artikulierten sie die Zungenlaute oft lateral.

Am meisten aber war ich in meiner Ansicht von der Gleichheit der Kau- und der Sprechbewegungen durch die *ausnahmslose* Erfahrung beeinflusst, daß Stimmgestörte, mit der Kaumethode behandelt, *keinerlei* Schwierigkeiten hatten, von dem ursprünglich «sinnlosen» lauten Kauen, bei dem nicht Wörter, sondern sinnlose Silben zutage kamen, in das laute Kauen der Muttersprache überzugehen. Mit andern Worten, daß sie die in der Muttersprache gebräuchliche Artikulation ohne Schwierigkeit mit der Stimme des primitiven Kauens verbinden konnten.

Hier muß besonders hervorgehoben werden, daß bei Anwendung der Kaumethode der Patient nicht nur nicht aufgefordert wird, die beim «sinnlosen» Kauen aufgetauchte Stimme nachzuahmen, sondern daß er davor gewarnt wird, *irgendeine* Stimme zu beabsichtigen. Die Stimme, die beim Kauen herauskommt, ist die von uns gewünschte. Aber diese Stimme ist – nach unsern genauen Beobachtungen – dieselbe, die beim «sinnlosen» Kauen erschien!

Es scheint mir fraglich, ob eine gemeinsame Wurzel von lautem Kauen und Sprechen, hätten sich beide Funktionen seit Jahrtausenden voneinander getrennt, sofort wieder aktiv werden könnte! Das ändert nichts daran, daß, wie *Trojan* dargetan hat, kortikale und extrapyramidale Bahnen und Zentren an der Sprache (und wohl auch am Kauen) beteiligt sind. Ich gebe zu, daß der Grad der Beteiligung beim Kauen und beim Sprechen, besonders neu zu erlernender Laute, verschieden ist. Aber beide Systeme sind wohl stets beteiligt.

Ferner möchte ich darauf hinweisen, daß die meisten Menschen, wenn aufgefordert, *ohne vorausgehende Übung* auf einer Seite des Mundes sprechen können (wie es ausnahmsweise geschieht, wenn man einem Nachbarn etwas sagt, das andere nicht bemerken sollen). Ich erinnere mich, daß ich als Kind mit andern Kindern ein solches Spiel vollführte.

Es ist merkwürdig, daß sich die Sprache, von den erwähnten Ausnahmen abgesehen, generell nach vorne bewegt und nicht nach

der Seite, wo man doch oft zu Leuten spricht, die rechts oder links vom Sprecher sitzen oder gehen, vermutlich, weil der Atem außerhalb des Sprechens nach vorne geht. Immerhin kann sich das Sprechen sofort nach einer Seite richten, wenn man es will!

Ich bin den drei Opponenten dankbar dafür, daß sie mich veranlaßt haben, die Kautheorie neuerdings zu überprüfen. Endlich sei noch *H. Beebes*¹⁶ Probe für die Identität von Kau- und Sprechbewegungen erwähnt: Die Versuchsperson wird angewiesen, ein Stückchen Backwerk im Munde zu halten und irgend etwas zu sagen, ohne die Backware zu kauen. Dann wird ihr gezeigt, daß durch das Sprechen die Backware zerkaut worden ist.

Zusammenfassung

Der Annahme von *Stein, Orthmann* und *Trojan*, daß sich die Sprechbewegungen aus den Kaubewegungen entwickelt haben, aber nicht mit ihnen gleichartig sind, besonders weil Kauen (aber nicht ausschließlich!) lateral erfolgt, werden folgende Beobachtungen entgegengehalten. 1. Negersprachen enthalten zahlreiche laterale Klicks, die sich leicht mit Vokalen verbinden. 2. Laterale Sigmatismen sind sehr häufig, laterale L-Laute kommen vor, manche vorher hörstumme Kinder sprechen oft anfangs mit lateraler Artikulation der Zungenlaute. 3. Während der Anwendung der Kaumethode zur Heilung von Stimmstörungen besteht niemals eine Schwierigkeit, vom «sinnlosen» Kauen zum «Kauen» beim Sprechen der Sprache überzugehen, obwohl jeder Patient davor gewarnt wird, etwa die Stimme, die während des «sinnlosen» Kauens herauskam, nachzuahmen. Er soll einfach wieder beim Sprechen die Stimme herauskommen lassen, die bei dem Gedanken entsteht, daß er jetzt eben auch nichts anderes tut als laut kauen. Die Art der Stimme ist stets dieselbe bei «sinnlosem» Kauen und beim Kauen während des Gebrauchs der Muttersprache. 4. Die meisten Menschen empfinden keinerlei Schwierigkeit, wenn sie der Aufforderung nachkommen, durch einen der beiden Mundwinkel zu sprechen, als ob sie jemand, der auf einer Seite des Sprechenden steht, ein Geheimnis sagen würden.

Aus all dem ist der Schluß zu ziehen, daß Stimme und Mundbewegungen beim Sprechen und lauten Kauen gleichartig sind.

Summary

The following observations are meant to contradict the opinion of *Stein, Orthmann* and *Trojan* that the speech movements have developed from the chewing movements but are not identical, especially because chewing proceeds (although not without exception) laterally. 1. Some negro languages contain numerous lateral clicks which can easily be connected with vowels. 2. Lateral lisps are very frequent, lateral L sounds exist and some originally hearing-mute children sometimes start talking with lateral articulation of tongue sounds. 3. During the use of the chewing method in the treatment of voice troubles there is never any difficulty in proceeding from "nonsense" chewing to chewing meaningful language, although the patient is always warned against imitating the voice which was evoked with nonsense chewing. The patient is asked just to let the voice come out which develops by the thought that he does not do anything but chew aloud. When the method is properly applied the voice is always the same in the mother tongue as in the "nonsense" chewing. 4. Almost no one encounters any difficulty in speaking out of one of either angle of the mouth, as if speaking secretly to a person located to the right or left.

From all these points, the conclusion is made again that the voice and the mouth movements in talking and in chewing are identical.

Résumé

A l'hypothèse de *Stein, Orthmann* et *Trojan* suivant laquelle les mouvements de la parole se sont développés à partir des mouvements de mastication, mais sans leur être identiques, en particulier parce que la mastication est latérale (quoiqu'elle ne le soit pas exclusivement), on peut faire les objections suivantes:

1^o certaines langues africaines contiennent des clics latéraux qui se combinent facilement avec les voyelles.

2^o les sigmatismes latéraux sont très fréquents, il y a beaucoup de L latéraux; beaucoup d'enfants sourds commencent à parler avec une articulation latérale des sons linguaux.

3^o Dans l'application de la méthode de la mastication au traitement des troubles de la parole, on ne rencontre jamais de difficultés pour passer de la simple mastication à la mastication avec emploi de la langue, même si on demande au patient d'essayer d'imiter le bruit qu'il faisait pendant la simple mastication. En parlant il doit simplement laisser se reproduire ce bruit, croyant qu'il ne fait encore rien d'autre que mastiquer tout haut... Le type de bruit produit par simple mastication est identique au bruit produit par la mastication avec emploi de la langue maternelle.

4^o La plupart des individus n'éprouvent aucune difficulté, si on les y invite, à parler du coin de la bouche, comme s'ils disaient un secret à une personne se trouvant à leur côté.

Il faut conclure de tout cela que le «bruit» et les mouvements de la bouche sont identiques quand on parle et quand on mastique tout haut.

Literaturverzeichnis

1. *Beebe, H. H.*: Practical Aspects of Chewing Therapy. *Folia Phoniatr.* 8: 107-117 (1956).
2. *Brodnitz, S. F.*: Keep Your Voice Healthy (New York 1953).
3. *Cabanas, R.*: Generalization of the Chewing Method in Logopedics and Phoniatriy. *Folia Phoniatr.* 4: 249 (1952).

4. *van Dantzig, B.*: Einzel- und Gruppenbehandlungsmethoden in der Sprachheilkunde. Mschr. Ohrenheilk. 1937: 610.
5. *Krech, H.*: Zur kombiniert-psychologischen Behandlung psychogener Stimmstörungen. Folia Phoniatr. 6: 120 (1954).
6. *Mohr, E.*: Die Therapie des Atemessens bei Stimmstörungen. Mschr. Ohrenheilk. 1935: 1197.
7. *Moolenaar-Bijl, A.*: Adem-Eten En Klank-Eten. Logopaed. En Phoniatr. 1954: 161
8. *Ringer, M.*: Zur Therapie des Atemessens bei Stimmstörungen. Mschr. Ohrenheilk. 69: 1200 (1935).
9. *Schmitz, F.*: Atemessen bei Trockenheitsempfindung organischen und psychogenen Ursprungs. Mschr. Ohrenheilk. 1937: 1401.
10. *Weiss, D. A. and Beebe, H. H.*: The Chewing Approach in Speech and Voice Therapy (mit Beiträgen von sieben Sprach- und Stimmtherapeuten) (New York 1952).
11. *Stein, L.*: Speech and Voice, p. 212 (London 1942).
12. *Orthmann, W.*: Sprechkundliche Behandlung funktioneller Stimmstörungen. Anwendung der Kaumethode für hyperkinetische Stimmstörungen, S. 15 (Halle a.d.S. 1956).
13. *Id.*: Grundlagen der Chewing Method (*Froeschels*) und ihre Anwendung bei stimmlichen Hyperkinesien. Spracharbeit 3: 49 (1957).
14. *Trojan, F.*: Zur Grundlegung der «Entwicklungsphonetik». Folia Phoniatr. 7: 99 (1955).
15. *Id.*: Zeichen, Silbe und Laut in entwicklungsgeschichtlicher Sicht. Phonetica 1: 63–81 (1957).
16. *Beebe, H. H.*: Practical Aspect of Chewing Therapy. Folia Phoniatr. 8: 107 (1956).

Adresse des Autors: Emil Froeschels, M. D., 133 East, 58th Street, New York 22, N.Y. (USA)

Delattre, P.: *Phonetica* 2: 226-251 (1958).

Université du Colorado, Boulder, Colo., et Laboratoires Haskins, New York, N. Y.

Les indices acoustiques de la parole: Premier rapport

Par PIERRE DELATTRE

(Fin)

Les résultats

Bibliographie: Nous citerons les travaux qui, sauf erreur ou omission de notre part, ont contribué à la connaissance des indices acoustiques de la parole (serait-ce même par le stimulant de conclusions erronées) au cours des dix années passées – environ 1947 à 1957. Les numéros suivent l'ordre chronologique.

La première question qui se pose en regardant un spectrogramme est: «Quels sont, dans ces 8000 cps de formants, dans cette richesse de traits acoustiques, les traits pertinents, du point de vue linguistique?» Dès que le premier synthétiseur des laboratoires *Haskins* a été au point, que les tests de parole artificielle résultant de la reconversion de spectrogrammes «naturels» ont été satisfaisants, le travail de défrichage a visé à répondre à cette question. Les nombreux formants ont été successivement couverts (un à un, puis en groupes) et les résultats de ces omissions ont été soumis à l'identification linguistique par l'oreille ^{11, 12, 13, 18}. Il est vite ressorti de cela qu'en dehors de quelques sons turbulents – surtout les frictions et explosions de dentales et alvéolaires – les trois formants les plus bas, souvent même les deux formants les plus bas, comprenaient tous les principaux traits pertinents. Partis de là, on a cherché jusqu'à quel point la peinture de spectrogrammes au pinceau permettait de simplifier l'aspect visuel, encore assez complexe, des deux ou trois premiers (plus bas) formants du spectre naturel, sans perdre l'intelligibilité du son issu de la machine. On a aussi poussé la simplification jusqu'à la perte partielle de l'intelligibilité. Ainsi, en remplaçant successivement les sinuosités de chacun des trois

formants pertinents par des lignes droites, on a simulé, pour l'oreille, l'effet d'un raidissement de mâchoire (F1), ou l'effet d'un raidissement de la langue (F2). En ralentissant la machine, ou inversement (ce qui ne change pas la fondamentale), on a fait partiellement changer les modes d'articulation: telle consonne sourde devenait sonore, telle explosive devenait fricative ou semi-voyelle, telle liquide ou semi-voyelle devenait voyelle ou diphtongue. En variant au pinceau la fréquence des ondes turbulentes et la direction des transitions de F2 et F3, on a surtout fait changer les lieux d'articulation. En général, en manipulant tous les aspects imaginables du spectrogramme, on a pu isoler des indices acoustiques et percevoir les limites dans lesquelles on peut faire varier leurs dimensions pour spécifier leurs rôles individuels dans la perception.

Explosives, ou Occlusives Orales

En tant que classe de consonnes, les occlusives se distinguent surtout par le degré d'interruption du son buccal (tenue), par la brièveté du son turbulent intense (explosion) qui suit, et par la rapidité des transitions qui mènent à la voyelle suivante ou qui viennent de la voyelle précédente. C'est la classe qui a été le plus étudiée, probablement parce qu'elle a paru la plus provocante. (Les fricatives ont paru si simples à synthétiser qu'on s'en est peu soucié au début.)

Les explosions. La première expérience systématique, à l'aide de SP, a été organisée pour étudier les effets des explosions d'occlusives initiales sourdes¹⁶. Les syllabes à identifier sont composées d'une explosion synthétique suivie d'une voyelle synthétique à deux formants droits de trois harmoniques par formant. Les explosions prennent la forme visuelle d'ovales verticaux de 600 cps et 15 ms. On leur donne 12 fréquences différentes (de 360 à 4320 cps) qui se combinent tour à tour avec chacune des 7 voyelles cardinales [i e ε a ɔ o u] pour un total de 84 syllabes synthétiques. On a fait entendre ces syllabes, enregistrées sur ruban magnétique en ordre de hasard, à 30 sujets non phonéticiens en leur demandant de les identifier comme /p/, /t/ ou /k/. Les résultats sont clairs: Les explosions hautes, en dessus de 3000 cps environ, sont identifiées comme /t/, les autres, en dessous de 3000 cps, comme /k/ ou /p/ selon qu'elles sont situées juste au dessus du début de F2 (/k/) ou ailleurs (/p/). L'examen des résultats révèle aussi que l'effet de la fréquence

de l'explosion n'est pas indépendant de la voyelle: dans un cas tout spécialement frappant, une même explosion, d'une fréquence de 1440 cps, est entendu comme /p/ quand elle est unie à [i] et comme /k/ quand elle est unie à [a]. D'autre part des explosions de fréquences extrêmement différentes se font entendre comme la même consonne. Donc, d'une part, un même son peut s'identifier de deux manières différentes; d'autre part, deux sons fort différents peuvent s'identifier de la même manière. Deux hypothèses se forment déjà, qui seront confirmées dans des expériences ultérieures. a) Dans la parole, la plus petite unité *acoustique* est la syllabe. b) S'il existe un «invariant» qui permette de distinguer un lieu d'articulation consonantique d'un autre, il est plutôt dans le geste articulatoire que dans le trait acoustique: la forme acoustique de la parole serait perçue, non directement, mais indirectement par référence au geste articulatoire qui est le même pour plusieurs valeurs acoustiques différentes.

Les explosions ont été étudiées par la synthèse dans deux autres travaux ^{51, 52}. Les explosions y sont jointes, non à des formants droits (voyelles à l'état stable), comme dans l'expérience précédente, mais à des formants commençant par des courbes de transition consonne-voyelle telles qu'on en voit sur les spectrogrammes.

Dans ⁵², on a une vaste étude de toutes les combinaisons appropriées de trois variables: transitions de F2, transitions de F3, et explosions. Seule la voyelle américaine [æ] est combinée à ces trois variables. En ce qui concerne les explosions, nous trouvons là 294 modèles différents (patterns) de syllabes synthétiques identifiées par 26 sujets comme /b/, /d/, ou /g/: 7 fréquences d'explosion jointes à 7 courbes de transition de F2 ($7 \times 7 = 49$ syllabes), puis jointes aux 35 combinaisons de 5 transitions de F3 avec 7 transitions de F2 ($35 \times 7 = 245$). (Une transition fixe de F1 est toujours présente sous une forme qui rend les syllabes sonores.) Ces 294 modèles de syllabes avec explosion peuvent se comparer avec les modèles sans explosion. Les résultats sont entièrement d'accord avec l'expérience de ¹⁶ mais vont plus loin: les explosions de haute fréquence favorisent les jugements de /d/, celles de basse fréquence, sauf la plus basse des 7, les /g/, d'abord (en descendant) aux dépens des /d/, puis aux dépens des /b/. Les meilleurs /g/ ont l'explosion juste au dessus de la transition (cf. ¹⁶). Enfin la fréquence la plus basse, qui ne favorise ni les /d/ ni les /g/, ne favorise que très peu les /b/. De cela il ne faut pas conclure que la perception du lieu d'articu-

lation labial dépend seulement des transitions – nous savons qu'en l'absence des transitions, certaines fréquences d'explosion font fort bien percevoir le lieu d'articulation labial¹⁶. On doit simplement conclure que, dans la perception du lieu d'articulation labial, le rôle des explosions est sans doute bien moins important que celui des transitions.

Au total les effets des explosions sont faibles comparés à ceux des transitions, pour les trois consonnes, et cela malgré le fait que les explosions de ces expériences sont probablement plus concentrées en fréquence que dans la parole naturelle. Mais n'oublions pas que seule la voyelle [æ] a été employée. Avec une voyelle arrondie, le rôle des explosions dans la perception du lieu d'articulation serait probablement beaucoup plus fort.

On trouve quelques données sur les explosions, du point de vue du mode d'articulation, dans une étude sur les affriquées⁵¹. L'une des distinctions entre la classe des affriquées et la classe des explosives étant dans la durée du son turbulent, il semble que la consonne est identifiée comme explosive (et non plus affriquée) à partir d'une durée maxima de 30 ms. Cette étude est faite par recollage aussi bien que par synthèse.

La première expérience importante par recollage¹⁹ a voulu vérifier les résultats de ¹⁶ dans la parole naturelle. La vérification a été positive. Les bandes magnétiques des syllabes [ki], [ka], [ku], ont été coupées juste après l'explosion, puis les portions coupées ont été recollées à des voyelles [i], [a], [u], sans transitions. Entre autres résultats, l'explosion de [ka], jointe à [i], est identifiée comme /pi/ par 93 pour cent des sujets, et jointe à [u], comme /pu/ par 99 pour cent des sujets. Nous avons donc, de même que dans l'expérience synthétique¹⁶, une même explosion entendue comme /k/ ou /p/ selon qu'elle est unie à [a] ou à [i].

On trouve d'autres recollages d'explosives (et autres consonnes) dans une étude qui met littéralement à l'épreuve le principe de commutation³⁶, et les résultats sont du même genre que dans ¹⁶. Chaque fois que la voyelle qui suit un élément consonantique fixe est changée, la perception de la consonne change aussi. D'ailleurs les résultats de toute commutation par recollage sont prévisibles d'après ce que l'on sait maintenant sur les transitions.

Dans une analyse détaillée des explosives sourdes et sonores du danois²³, on trouve toute la complexité des données que fournit typiquement l'analyse au sujet des intensités, durées, et diverses

concentrations d'énergie sur l'échelle des fréquences, pour les explosives /p t k b d g/ devant toutes les voyelles danoises. Combien de ces traits sont distinctifs? Seule l'épreuve de la reconversion synthétique en son pourrait le déterminer. (Ainsi le fait que l'explosion du /p/ n'est pas concentrée en fréquence mais s'étend sur presque toute l'étendue en fréquence du spectre ne veut pas nécessairement dire qu'elle n'a pas de rôle dans la distinction des lieux d'articulation. La synthèse pourrait montrer que certaines portions – différentes selon les voyelles – de cette haute étendue de bruit en fréquence jouent un tel rôle.) En général, les hypothèses, présentées dans la conclusion de cette étude, sur le rôle des explosions et transitions dans la perception des occlusives, ne sont pas d'accord avec les résultats ultérieurs obtenus par la synthèse ^{21, 26, 52}. Mais notons aussi que les «spéculations» de ^{11, 13}, elles non plus, ne sont pas d'accord avec les résultats ultérieurs de ^{21, 26, 52}. A cette époque, ni le principe du Locus de F2, ni le rôle des transitions de F3 n'étaient encore connues.

Deux études, qui comparent les occlusives finales, quand elles sont privées de leur détente et quand elles ne le sont pas, arrivent à des résultats comparables. Dans l'une³⁷, l'omission de la détente est simulée par le sujet, qui enregistre sans rouvrir la bouche; dans l'autre⁴⁵, la syllabe est d'abord enregistrée avec détente, puis la détente est coupée. (Le lecteur se rend compte qu'en coupant la détente d'une consonne finale, on la prive de l'explosion, ainsi que des embryons de transition qui peuvent suivre l'explosion, et qu'il ne reste, pour percevoir le lieu d'articulation, que les transitions implosives qui précèdent la tenue.) Les résultats les plus intéressants sont ceux qui montrent, dans les deux études, que les consonnes qui souffrent le plus de l'absence d'explosion sont /k g/ devant [u] (devant les voyelles postérieures arrondies, en général). Ce résultat indique que la perception du lieu d'articulation de /k g/ devant voyelles postérieures arrondies dépend beaucoup de l'explosion et peu des transitions. D'ailleurs le fait était à prévoir: pour /k g/, la transition de F2 devant [o u] se dirige, non pas vers le Locus vélaire, mais vers le Locus labial; une fois [uk] privé de son explosion, c'est [up] ou [u] qui devrait s'entendre – et c'est ce qui arrive dans les tests perceptuels de ^{37, 45}. Dans les tests de ³⁷, c'est, en plus, après /l/ sombre et /r/ sombre que /k/, privé d'explosion, est mal perçu; or les formants 1 et 2 de /l/ et /r/ sombres sont tout proches de ceux de [o].

L'étude⁴⁵ examine encore les explosions par l'analyse et le filtrage. On y établit (mais de façon incertaine) que les explosions isolées de leur contexte sont identifiables. Puis, sur la base des intensités-fréquences des explosions de /p t k b d g/ après et avant six voyelles représentatives des diverses positions articulatoires, on s'efforce de découvrir, par filtrage, deux paires de traits binaires qui permettraient leur identification à l'état isolé. Ces deux paires de traits distinctifs passeraient-ils l'épreuve de la synthèse? Quoi qu'il en soit, on trouve dans cet article des données spectrales précieuses sur les explosions, et en gros, ces données sont d'accord avec les résultats obtenus par la synthèse^{16, 52}: pour /t d/, les fréquences sont hautes; pour /p b/ elles sont basses; et pour /k g/, elles sont intermédiaires mais dans une très grande marge de fréquences parce qu'elles suivent les transitions de F2, qui varient d'environ 3000 cps à 600 cps.

Les transitions d'occlusives. Jusqu'ici, on n'a trouvé d'indices acoustiques que dans les trois premiers formants. Pour abrégé, appelons leurs transitions T1, T2, et T3. Les indices trouvés pour T2 et T3 se rapportent presque entièrement au lieu d'articulation (comme la fréquence des explosions). Les indices trouvés pour T1, au contraire, se rapportent aux modes d'articulation: distinction entre classes de consonnes; distinction entre sourdes et sonores.

T1: Très tôt, il a été observé, par l'analyse des spectrogrammes, que F1 est d'autant plus haut (fréquence) que les voies buccales sont plus ouvertes^{2, 8}. Appliqué aux consonnes, cela indiquerait que plus la consonne sonore est ouverte, plus T1 devrait commencer haut. Mais aucune investigation systématique de cette corrélation n'a été faite pour les consonnes, et il nous faudra réunir des idées éparses dans les diverses études.

Dans les recherches sur T2^{21, 52} et T3⁵², pour obtenir des occlusives sonores, il a fallu que T1, assez rapide, commence aussi bas que possible (nous ne savons pas si cela correspond effectivement à zéro cps ou à 120 cps – la fondamentale de SP). Dans²¹, pour obtenir des occlusives *nasales*, on a dû faire partir T1, semble-t-il, de la fréquence de FN1 (le plus bas formant de tenue nasale – vers 250 cps) et joindre verticalement le point de départ à la voyelle contiguë, ce qui fait paraître, visuellement, que T1 est droit et part du niveau même de la voyelle contiguë. Dans l'étude du Locus de F1 pour les explosives²⁶, les variations de fréquence de F1 *droit*, combinées à F2 courbé, indiquent que le point de départ le plus bas,

pour T1, est le meilleur pour les explosives, et que, à mesure que ce point monte en fréquence, on se rapproche de la perception des classes de consonnes plus ouvertes. L'examen des spectrogrammes de fricatives indique, en général, pour T1, un départ moins bas que chez les explosives. Quant aux liquides et semi-voyelles initiales, on trouve, dans ⁴⁹, que leurs T1 doivent partir d'assez haut – près de 400 cps en moyenne – si l'on veut éviter toute perception d'occlusive.

On voit qu'il reste beaucoup à faire pour préciser le rôle de T1 dans la distinction des classes de consonnes.

La vitesse de transition et la durée de T1 contribuent aussi à des distinctions de classe. Ces deux facteurs, variés à la fois pour T1 et T2³³, ont permis de distinguer entre les trois classes suivantes: voyelles, semi-voyelles, explosives sonores. Par changement de durée-vitesse de T1-T2, /u/ est passé à /w/, puis à /b/; /i/ est passé à /j/, puis à /g/; et si l'on avait eu des sujets français pour faire les identifications on aurait sans doute trouvé que, par les variations des mêmes facteurs, /y/ peut passer à /ɥ/, puis à /d/. Le changement de semi-voyelle à consonne est plus net que celui de semi-voyelle à voyelle. Le changement de /b/ à /w/ se fait quand la transition a une durée d'environ 40 ms; celle de /g/ à /j/, 50 à 60 ms.

La forme *implosive* de T1-T2, et leur forme *explosive*, sont présentées comme contribuant à la perception du point de coupe syllabique (respectivement *après* ou *avant* la consonne) dans ³⁵. L'investigation est faite par synthèse.

Enfin nous verrons plus loin que certaines dimensions de T1 semblent contribuer à la distinction entre /p t k/ et /b d g/ communément appelée sourde-sonore³².

T2. Pour une durée assez courte (ou une vitesse assez rapide), les T2 sont sans doute les plus puissants indices de distinction entre les lieux d'articulation. Excepté chez /k/ devant voyelle arrondie, ils sont plus effectifs que les explosions, ce qui est compréhensible, car ils ont, comme les formants vocaliques, une beaucoup plus grande audibilité que les bruits sourds d'explosion. La durée de T2, chez les explosives, est en moyenne de 50 ms, mais elle tend à être plus courte que cela chez les labiales et plus longue chez les dentales devant voyelles postérieures.

Les dimensions de T2 qui contribuent à l'identification du lieu d'articulation sont a) sa direction, dite positive si elle atteint plus haut que F2 de la voyelle, et négative si elle atteint plus bas;

b) la différence de fréquence entre son début et le moment où elle rejoint F2 de la voyelle (cette dimension est généralement donnée, dans les travaux de *Haskins*, par un multiple de 120 cps – ainsi une transition de –3 atteint une fréquence de 360 cps en dessous du formant correspondant de la voyelle.

Une vaste étude de T2 par la synthèse²¹ a suivi de peu l'étude¹⁶ des explosions. Elle comprenait 11 variations de T2, jointes chacune à 7 voyelles cardinales [i e a ɔ o u], et ceci répété pour les occlusives sourdes, sonores, et nasales, pour un total de 231 modèles de spectrogrammes artificiels, reconvertis en son, et identifiés par 33 sujets. Aucune explosion n'était employée dans ces modèles. La sonorité était obtenue en faisant partir T1 de zéro (ou 120 cps), la sourdité en supprimant le début de T1, et la nasalité en faisant partir T1 du niveau de F1 et en ajoutant trois formants nasals dans la tenue.

Les résultats, fort complexes, montrent un T2 différent, non seulement pour chaque lieu d'articulation, mais pour chaque voyelle combinée à chaque lieu d'articulation. Par ailleurs, les résultats pour nasales sont fort semblables à ceux des sonores et des sourdes.

Locus. Dans la recherche d'un invariant par lieu d'articulation, on a remarqué que toutes les T2 perçues labiales convergeaient virtuellement vers une fréquence basse (quelle que soit la voyelle de la syllabe), toutes les T2 perçues dentales (ou alvéolaires) vers une fréquence intermédiaire, et toutes les T2 perçues vélaires (ou palato-vélaires) vers une fréquence haute. (Cela laissait une petite région d'ambiguïté car devant les voyelles postérieures arrondies [ɔ o u], aucune T2 n'était perçue clairement vélaire – problème qui a été résolu depuis.) On a donné le nom de Locus à ce point de convergence virtuel des transitions qui ont perceptuellement un même lieu d'articulation.

La spécification, en fréquence, du Locus a fait l'objet d'une longue recherche par la synthèse. Le Locus corrélatif à chaque lieu d'articulation a été déterminé, non par extrapolation de courbes de T2, mais par variation de formants droits, évitant ainsi l'erreur que pourraient occasionner les courbes: en faisant varier, du haut en bas de l'échelle des fréquences un F2 droit (T2 zéro) combiné à un T1 fixe, courbé à souhait pour produire une explosive sonore, on a obtenu un /g/ quand F2 droit était à 3000 cps; puis, en abaissant la fréquence de F2 droit, le /g/ s'est perdu et le /d/ a commencé à s'entendre pour arriver à son maximum de perceptibilité à

1800 cps; en continuant à abaisser la fréquence de F2 droit, le /d/ s'est perdu et le /b/ s'est fait entendre pour arriver à son maximum vers 700 cps. Ensuite il a fallu déterminer la durée qui sépare l'extrémité des T2 de leurs Locus respectifs. Cela s'est fait par des coupes successives de transitions partant du Locus même. On est arrivé à la durée moyenne de 50 ms. Le Locus, ainsi spécifié, fournit un invariant pratique par lieu d'articulation. Il permet de définir une transition d'occlusive sans référence à la voyelle de la syllabe. Toute T2 d'occlusive peut se décrire comme ayant une durée d'environ 50 ms et se dirigeant vers le Locus du lieu d'articulation qu'elle fait percevoir par une ligne virtuelle qui l'atteindrait en 50 ms. Les différences de fréquence entre les extrémités réelles des transitions qui sont perçues par un même lieu d'articulation sont évidemment dues à l'anticipation articulatoire de la voyelle contiguë à la consonne. (La corrélation articulatoire des trois Locus, ainsi que la non application du Locus vélaire aux voyelles arrondies, ont été clairement établies sur rayons-X cinématographiés, mais nous n'avons pas ici la place de traiter les corrélations physiologiques des indices acoustiques.)

Des recherches en cours indiquent que les occlusives vélares devant voyelles arrondies, *dans la parole naturelle*, ont leur principal indice acoustique de lieu d'articulation dans la fréquence de l'explosion. Si, dans la parole naturelle, T2 d'une syllabe comme [go] ne se dirige pas vers le haut Locus vélaire, c'est à cause de l'arrondissement qui maintient très basse la fréquence du début de la transition. Mais dans la parole artificielle, on peut obtenir un [go] sans peindre d'explosion, seulement par des T2 positives, dirigées vers le Locus vélaire de 3000 cps et assez longues pour dépasser le niveau du Locus dental de 1800 cps.

Dans ²² on trouvera une présentation du concept du Locus tel qu'il a été compris personnellement par un visiteur aux laboratoires *Haskins*. Mais disons ici que l'hypothèse donnant au Locus la fréquence du résonateur buccal avant l'explosion de la consonne n'a plus cours.

Un Analogue électrique des cavités buccales a essayé de vérifier le concept du Locus ⁴². Cet analogue simule trois variables articulatoires au moyen desquels il produit synthétiquement des sons soutenus (genre voyelles): le point de constriction linguale, le degré de constriction linguale, et le degré et la longueur de constriction labiale. Il ne produit donc pas de consonnes, mais on peut quand

même l'utiliser pour l'étude des consonnes en observant sur les spectrogrammes successifs les effets de chaque ajustement de variable. Les résultats ainsi obtenus pour les transitions de /b d g/ sont à peu près d'accord avec les Locus respectifs, à condition de comprendre que, tout au cours de l'étude, le terme «Locus» a été confondu avec «début de transition». Rien d'extraordinaire à ce que les débuts de transition (et non les Locus) varient par anticipation de la voyelle – anticipation dont on a tenu compte mais qu'on a supposée (dans ⁴²) bien plus marquée qu'elle ne l'est, surtout pour /b/, d'après les radiogrammes cinématographiques de l'articulation de /b d g/.

Notons, avant de quitter le sujet, que la notion de Locus ne s'appliquera pas seulement aux transitions d'occlusives, mais peut-être aux transitions de toutes les consonnes. Il semble bien que les /f/ ont le même Locus que les /p/, les /s/ que les /t/, etc., ou en tout cas à peu de chose près.

Dans ⁵², nommé déjà pour les explosions, les variations de T2 sont étudiées très soigneusement, ainsi que leurs combinaisons avec des variations de T3, ou d'explosions, ou des deux. Devant la voyelle [æ], et T1 fixe étant peint de façon à produire des occlusives sonores, 7 variations de T2 sont étudiées: -6, -4, -2, 0, +2, +4, +6. Ce sont les mêmes dimensions que dans ²¹ mais avec omission des transitions impaires, pour simplifier. Les résultats confirment entièrement ceux de ²¹. Les seules transitions des deux premiers formants (ni explosions, ni T3) suffisent à distinguer /b d g/ entre eux. C'est /b/ qui dépend le plus de T2 et /d/ qui en dépend le moins – /d/ dépend de T3 bien plus que les deux autres. Les jugements de /b/ sont à peu près de 100 pour cent à -6, -4, et -2, puis diminuent brusquement. A zéro, /d/ atteint presque 90 pc et à +2, presque 100 pc, puis /d/ diminue brusquement et fait place à /g/, qui atteint 95 pc à +4, et 100 pc à +6.

Les résultats de l'étude de T2 *par l'analyse* concordent parfaitement avec ceux de la synthèse, mais naturellement ils sont plus vagues – c'est précisément parce que les spectrogrammes sont difficiles à lire, surtout dans les transitions, que la synthèse rend de tels services.

Quatre études de T2 par l'analyse sont à noter.

Les remarques sur T2 abondent dans ¹, où la notion du «hub» peut être considérée comme un avant-poste de celle du Locus, bien qu'en réalité les deux notions diffèrent considérablement.

Les analyses de T2 dans ³, déjà mentionnées dans l'introduction, prévoyaient avec perspicacité le rôle que la synthèse allait confirmer et préciser.

On trouve dans ²³ de bonnes analyses de T2 dans les explosives danoises, qui indiquerait que le Locus des labiales danoises est moins bas que celui des labiales anglo-américaines et latines.

Enfin, les analyses de T2 dans ²³ confirment entièrement les résultats de ²¹, ²⁶, ⁵² obtenus par la synthèse.

T3. Rien n'a encore été publié qui porte spécialement sur T3, mais nous pouvons dire que les résultats d'une étude détaillée, en cours de publication, sont d'accord avec ceux de ⁵² que nous résumons ci-dessous en notant pourtant qu'ils ne s'appliquent qu'à la voyelle [æ].

Le problème de T3 est infiniment plus simple que celui de T2 puisque F3 a à peu près la même fréquence pour toutes les voyelles (il est un peu plus haut pour un [i] bien cardinal). En gros on peut dire que T3 est positive pour les dentales, et négative pour les labiales et les vélaires. Parmi les T3 négatives, toutes contribuent aux labiales plus qu'aux vélaires, et les moins basses contribuent plus aux vélaires que les plus basses. La perception du lieu d'articulation dental doit beaucoup à T3 (avec certaines voyelles, peut-être plus qu'à T2); celle du lieu d'articulation labial, moins; et celle du lieu d'articulation vélaire encore moins (pour les labiales, c'est T2 qui domine, et pour les vélaires c'est soit T2, soit l'explosion).

Dans ⁵², où d'une part 5 variations de T3 (-4, -2, 0, +2, +4), d'autre part 7 fréquences d'explosions, sont combinées à chacune des 7 variations de T2, on a l'occasion de comparer les effets de T3 à ceux des explosions. En général, la contribution de T3 est nettement plus grande que celle des explosions pour /d/ et /b/; pour /g/, c'est l'inverse: la contribution des explosions est la plus grande.

D'après les remarques analytiques de ²³, les T3 d'occlusives danoises diffèrent quelque peu de ce qui précède. Quant aux analyses de ⁴⁵, elles sont à peu près d'accord avec ce qui précède. (Mais on sait les difficultés qu'il y a à distinguer les courbes de T3 sur les spectrogrammes.)

Les fricatives

Il a été établi dans ⁵¹ que les fricatives, en tant que classe de consonnes, se distinguent en partie des affriquées et des explosives

par la durée du bruit (son turbulent) ainsi que par la rapidité avec laquelle l'intensité initiale de ce bruit croît. La durée du bruit est relativement longue, et la vitesse de croissance de l'intensité relativement lente, pour les fricatives (voir données aux Affriquées). Le rôle de la vitesse des transitions dans la distinction de classe n'a pas été étudié systématiquement. Il est certain, par exemple, qu'entre les transitions rapides de /b/ et les transitions lentes de /w/, il existe un régime de transitions qui correspond à /v/, et il faudra déterminer là les rôles respectifs de T1, T2, et T3. Dans cette même distinction des fricatives comme classe, le rôle de la fréquence du début de T1 mérite aussi d'être étudié.

Quant aux indices qui permettent de distinguer entre les diverses fricatives, rien n'ayant paru, nous ne donnerons que de vagues indications, hypothétiques, en partie basées sur une communication dont l'abstrait est dans *JAS.* 26: 952. D'après la synthèse, ces indices se trouvent dans les transitions supérieures (T2, T3), et dans les bruits de friction.

Les rôles de T2 et T3 dans la perception du lieu d'articulation sont sans doute comparables à ceux des occlusives, mais l'étude systématique n'en a pas été faite. T2 et T3 devraient donc pouvoir se décrire par les Locus corrélatifs aux lieux d'articulation, mais il faut s'attendre à ce qu'elles jouent un rôle moins important dans les fricatives que dans les occlusives car les bruits de friction sont plus audibles que les bruits d'explosion.

Il faut peut-être diviser les fricatives en trois sous-classes qui se distingueraient entre elles par l'intensité de la friction, l'étendue en fréquence de la friction, et les transitions: /s ʃ/ auraient une forte intensité et une étendue moyenne; /θ f/ une faible intensité et une grande étendue (presque toute la fréquence du spectre sur les spectrogrammes Kay); /ç X/ une intensité moyenne et une étendue étroite. Le rôle des transitions ne doit pas être négligeable car, à la synthèse, une même friction (ambiguë vers 3500 cps) s'entend comme /s/ ou /ç/ selon qu'on la relie à la voyelle par une T3 (positive – dentale) ou par une T2 (positive – palato-vélaire).

A l'intérieur de ces trois classes, les distinctions sont simples: /s/ se distingue de /ʃ/ principalement par la fréquence de la friction (/s/ descend environ jusqu'à 3500 cps, /ʃ/ jusqu'à 2000); /θ/ se distingue de /f/ principalement par les transitions (ils ont à peu près, /θ/ le Locus dental, /f/ le Locus labial); /ç/ se distingue de /X/ par les deux – fréquence de friction, et transitions.

La fricative /h/ serait caractérisée par un bref son turbulent à la fréquence de F2 (et peut-être F3) de la voyelle contiguë – donc par l'absence de transitions et l'absence de F1. La glotte étant grande ouverte, la friction pertinente pour /h/ serait celle qui résonne dans la cavité qui est antérieure au point de constriction vocalique. Ceci, contrairement aux voyelles chuchotées, qui résonneraient dans toutes les cavités supérieures à la glotte, le point de constriction qui produit le son turbulent étant aux cordes vocales mêmes – c'est pourquoi elles posséderaient un F1, bien que sourdes.

Un travail par l'analyse et le filtrage est à citer ³¹. On y étudie les spectres de frictions isolées de /f s ʃ/ et des sonores correspondantes, en toutes positions et prononcées par des sujets variés. Les résultats confirment ce qui a été dit plus haut sur la distinction entre /s/ et /ʃ/ par la différence de fréquence de la friction. Ils apportent un facteur nouveau: /f/ aurait fréquemment – mais pas toujours – une concentration d'énergie très élevée, aux environs de 8000 cps. Ceci est à vérifier par le premier synthétiseur qui atteindra cette fréquence.

Les Affriquées

Dans l'étude ⁵¹, déjà mentionnée, on examine les affriquées du point de vue du mode d'articulation: ce qui les distingue des fricatives et des explosives dans le son turbulent, en dehors du fait que les affriquées ont, comme les explosives, une interruption (occlusion buccale complète) que n'ont pas les fricatives.

On trouve deux indices acoustiques: la durée du bruit, et la rapidité de croissance de l'intensité du bruit (mesurée par la durée de la période pendant laquelle l'intensité croît, au début du bruit – appelons-la: durée de croissance).

En gros, comparé au bruit des fricatives, le bruit des affriquées est (après interruption) plus court en durée totale et plus court en durée de croissance. Comparé aux explosives, le bruit des affriquées est plus long en durée totale. Pour des valeurs moyennes de durée de croissance, on perçoit des fricatives sourdes quand le bruit total dure au moins 110 ms, des affriquées sourdes quand le bruit total dure au moins 50 ms, et des explosives sourdes quand le bruit total dure au *plus* 30 ms.

Les indices de lieu d'articulation des affriquées n'ont pas été étudiés systématiquement, mais il est clair qu'ils se trouveront,

comme pour les fricatives et les explosives: d'abord dans les transitions (Locus) puis dans la fréquence des bruits de friction. Il est probable que l'intensité du bruit et l'étendue de sa fréquence entreront en jeu.

Les Occlusives Nasales

Nous classons les consonnes nasales /m n ŋ/ parmi les occlusives parce qu'elles partagent avec les explosives (occlusives orales) la forme (vitesse) et la direction des T2 et T3. De plus, la fermeture nasale extérieure n'empêche pas leur production; l'ouverture extérieure des narines, d'ailleurs très petite, n'est donc pas pertinente – ce qui est pertinent ce sont les occlusions buccales, et la communication des cavités nasales avec les cavités buccales par l'abaissement du voile du palais.

Les indices acoustiques de mode d'articulation, aussi bien que de lieu, ressortent assez bien de l'étude ²¹, faite par la synthèse.

Mode. Les occlusives nasales se distinguent des occlusives orales a) par la forme de T1, qui semble partir du niveau de FN1 (environ 250 cps) et passer verticalement à celui de la voyelle contiguë, au lieu de partir de zéro ou 120, comme pour les occlusives orales sonores; et b) par les formants nasals de la tenue – qui remplacent le silence complet des occlusives orales sourdes, ou le ton très bas des occlusives orales sonores (qui correspond à la fondamentale et parfois un peu au deuxième harmonique entendus à travers les parois buccales et pharyngiennes). Dans les expériences de ²¹, les formants de la tenue nasale étaient les mêmes pour les trois consonnes /m n ŋ/, des recherches exploratoires ayant indiqué qu'ils ne jouaient qu'un rôle très faible dans la distinction des lieux d'articulation. Ils étaient aux fréquences de 240 cps, 1020 cps et 2460 cps. Les deux plus hauts de ces trois formants nasals étaient d'intensité extrêmement faible (environ 15 db de moins que ceux d'une voyelle normale à la même fréquence) et contribuaient très peu à la nasalité de la consonne. Le premier formant nasal, à 240 cps, était seulement un peu plus faible que celui d'une voyelle normale (environ 6 db de moins) et avait un puissant effet perceptuel de nasalisation. Il apparaît donc que le mode nasal des consonnes dépend de la forme de T1, d'un formant de tenue dans les 250 cps, et des formes de T2 et T3 semblables à celles des explosives de même lieu d'articulation. (L'importance du formant nasal d'environ 250 cps a été signalée pour la première fois dans ²⁰.)

Lieu d'articulation. a) Le rôle des transitions, défini par les Locus corrélatifs aux lieux d'articulation, est le même que pour les explosives. Il est très fort mais il n'est pas unique. b) La fréquence des formants de tenue nasale supérieurs à celui de 250 cps joue aussi un rôle dans la perception du lieu d'articulation – faible, mais certain. Des recherches par synthèse, ultérieures à ²¹, indiquent que la perception du lieu d'articulation labial est favorisée par la présence dans la tenue d'un F2 faible entre 1000 et 1500 cps, et par l'absence, ou la faiblesse de F3; la perception des lieux d'articulation dentals et vélares, par la présence dans la tenue, d'un F3 aux environs de 2300 (en plus de F2). On n'a pas trouvé jusqu'ici d'indice clair de distinction entre dentale et vélaire dans les formants de tenue nasale.

On a beaucoup étudié le rôle des formants de tenue nasale par la technique du recollage. Dans ³⁹, on a inclus, non seulement des interversions de tenue nasale avec les transitions qui suivent pour les trois consonnes /m n ŋ/, mais encore des interversions de tenue nasale avec les transitions qui suivent l'explosion des explosives sonores orales /b d g/. Les résultats confirment que l'indice du lieu d'articulation est presque entièrement dans les transitions; les tenues de nasales ont un rôle presque négligeable à l'initiale, et un peu plus apparent, mais encore très faible, à la finale.

On s'est aussi servi de l'Analogue électrique de la bouche ⁵⁰ pour produire synthétiquement les sons de tenue des trois consonnes nasales /m n ŋ/. La distinction des lieux d'articulation a été perçue (81 pc, 61 pc, 62 pc respectivement pour /m n ŋ/) – mieux perçue que les segments de tenue des nasales humaines de ³⁹: (96 pc, 36 pc, 12 pc). Il faut dire que pour ⁵⁰, les jugements avaient été faits par 9 sujets entraînés; pour ³⁹ par 50 sujets non entraînés. L'analyse spectrale du son de tenue nasale produit par l'Analogue confirme l'importance de F2 pour distinguer /m/ des deux autres, et indique une possibilité de distinguer /n/ de /ŋ/ par un formant au dessus de 3000 cps.

Les Liquides et Semi-Voyelles

Les /w j r l/ américains ayant certaines similarités spectrales de résonance, qui reflètent leur degré d'aperture articulaire (en moyenne plus grand que chez les occlusives, les affriquées et les fricatives), on les a étudiés ensemble en position initiale ⁴⁹. Le /r/

dont il s'agit ici est une rétroflexe apicale et palatale continue (sans battements).

Mode. Ces quatre consonnes semblent avoir acoustiquement en commun, pour se distinguer des autres consonnes: a) Pendant la tenue, un F1 de fréquence relativement haute (près de 400 cps de moyenne), qui les distingue surtout des nasales dont le bas formant de tenue ne peut pas dépasser 250 cps. b) Pendant la tenue, des formants supérieurs à F1, d'intensité plus grande que ceux de la tenue nasale mais plus faibles que ceux des voyelles. c) Des transitions *en continuité* avec les formants de tenue (les transitions des nasales peuvent être en discontinuité avec les formants de tenue). d) Une lenteur relative des transitions (en moyenne environ 100 cps, tandis que les transitions d'occlusives ont en moyenne 50 cps).

Lieu. Ces quatre consonnes se distinguent les unes des autres par la fréquence des formants de tenue supérieurs à F1, et les transitions allant de la tenue consonantique à la tenue de la voyelle contiguë. Les transitions peuvent se décrire par un Locus qui serait leur point virtuel de convergence corrélatif à un même lieu d'articulation. a) /w/ se distingue de /r l/ et de /j/ par le Locus de T2, qui est bas pour /w/ (environ 700 cps), moyen pour /r l/ (environ 1100 pour /r/ palatal, 1300 pour /l/ alvéolaire) et haut pour /j/ (environ 2700). b) /r/ et /l/ se distinguent entre eux par le Locus de T3, qui est relativement bas pour /r/ (environ 1500 cps) et haut pour /l/ (environ 2500 cps). T3 n'a pas d'effet notable pour /w/ et /j/. On peut ajouter trois divergences mineures qui contribuent sans doute aussi à la distinction entre ces quatre consonnes: c) Une durée de transition de 100 ms est acceptable pour les quatre, mais une durée un peu plus courte favorise /l/ contre /r/, et une durée un peu plus longue favorise /r/ contre /l/. d) Les formants de tenue sont moins indispensables chez les semi-voyelles que chez les liquides; et leur durée moyenne est plus courte chez les semi-voyelles (30 ms) que chez les liquides (60 ms). e) /j/ est amélioré par de la friction à une fréquence appropriée, ce qui n'est pas le cas des trois autres consonnes. Cela rapproche /j/ de la classe des fricatives, classe à laquelle il appartient s'il n'est autre que la sonore de /ç/.

Les Consonnes Syllabiques

Aucune étude systématique n'a encore paru. Des travaux en cours, par la synthèse, étudient la réduction d'intensité de F2, et

la forme implosive des transitions qui précèdent, comme indices de modes d'articulation qui distingueront entre les consonnes /l r m n ŋ/ en position syllabique et les voyelles qui ont leurs formants presque aux mêmes fréquences: ainsi, [mæd-l] se distingue de [mædo] par ces deux indices.

Les mêmes travaux étudient les fréquences de formants comme indices acoustiques de distinction entre les diverses consonnes syllabiques.

Sourdes et Sonores

A mesure qu'on découvre de nouveaux facteurs acoustiques du dit «voisement» ou de la dite «sonorité» des consonnes, la présence de la fondamentale (premier harmonique du spectre de la vibration des cordes vocales), qui est à l'origine de ces termes, prend une plus petite place et on arrive prudemment à remplacer les termes «distinction sourde-sonore» par «distinction du type p-b!». D'après ²³, en effet, la distinction /p b/ ne peut pas dépendre de la présence ou absence de vibrations des cordes vocales puisqu'elles ne vibrent ni pour l'un ni pour l'autre. Nous conserverons cependant, ici, la terminologie «sourde-sonore» – arbitrairement, ou dans le sens perceptuel.

a) La marque de sonorité qui est généralement présente sur les spectrogrammes pendant la tenue des occlusives sonores est appelée dans ¹: «barre de voix» (voice bar). Avec filtrage large à 300 cps, c'est la représentation de la fondamentale, à laquelle il s'ajoute une plus ou moins forte intensité du second harmonique, selon les sujets. La contribution de cette barre de voix à la perception de la sonorité a été amplement confirmé par la synthèse: ainsi, dans ⁵², afin que les syllabes [ba da ga] soient bien perçues «sonores» on a fait précéder les transitions d'un trait au niveau du premier harmonique faisant entendre la fondamentale pendant 60 ms. Mais on peut, dans la synthèse, faire percevoir la sonorité par plusieurs autres facteurs, en l'absence de la fondamentale. On peut aussi le faire par recollage: dans ⁴⁸, quand on remplace la tenue du /b/ de «Ruby» (qui contient la fondamentale) par un bout vierge de bande magnétique de même durée, on entend toujours «Ruby» plutôt que «Rupée» (oreilles anglo-saxonnes et latines). Pour perdre la sonorité perceptuelle, il ne suffit pas d'omettre la fondamentale, il faut, en plus, faire d'autres changements – par exemple, allonger le silence de la tenue (voir plus bas à /g/).

b) Chez les fricatives – qui gardent une certaine ouverture buccale pendant la tenue – le voisement peut comprendre, en plus de la barre de voix, une voyelle neutre, de basse intensité, parallèlement à la friction. La contribution de ce facteur est confirmée par la synthèse.

c) La simple présence de T1 semble contribuer très fort à la perception de la sonorité, et inversement. Dans ²¹, déjà, les occlusives sourdes étaient produites synthétiquement avec T1 très réduit. Des travaux en cours étudient la suppression de T1 comme facteur de surdité en positions initiale, implosive intervocalique ou de détente, explosive intervocalique ou de détente.

d) L'aspiration, c'est-à-dire acoustiquement: l'omission de T1 et la présence de son turbulent (inharmonique) au lieu de son périodique (harmonique) dans les 50 ou 60 premiers ms de T2 et T3, contribue nettement à la perception de la surdité. Noter que si T1 aussi a du son turbulent, il n'y a plus guère d'effet de surdité: sans doute parce qu'on a alors dans les 60 premiers ms, non plus de l'aspiration, mais de la voyelle chuchotée. L'aspiration serait donc semblable à la consonne /h/, et, comme /h/, s'articulerait la glotte grande ouverte, seules les résonances des cavités antérieures à la constriction vocalique étant alors perceptuellement effectives; au contraire, la voyelle chuchotée ayant tous les formants de la voyelle non chuchotée, y compris F1, sa constriction serait à la glotte, et toutes les cavités antérieures à la glotte seraient perceptuellement effectives.

e) La vitesse de transition de T1 a été étudiée dans ³², où une durée de T1 de 20 ms, ou un peu moins fait percevoir les occlusives comme sourdes, et une durée de 50 ms ou un peu plus les fait percevoir comme sonores.

f) La présence de T3 contribue légèrement à la surdité. Ainsi, c'est parce que les modèles spectrographiques artificiels de /b d g/, employés pour étudier les variations de T2 dans ²¹, n'avaient pas de T3 qu'il n'a pas été nécessaire d'ajouter une fondamentale; inversement, c'est parce que les modèles employés dans ⁵² avaient des T3 qu'il a fallu ajouter une fondamentale.

g) La durée relative de la tenue des consonnes intervocaliques (ou finales avec détente vocalique) est un facteur très puissant. Dans ⁴⁸, on a étudié par recollage de parole naturelle, les variations de durée de l'interruption buccale (tenue) des explosives: «Ruby» a passé à «Rupée» entre 60 et 100 ms de tenue, sans fondamentale,

et entre 80 et 120 ms de tenue, avec fondamentale. Dans ²⁷, on peut voir, entre autres choses, l'effet de la durée de la friction (tenue des fricatives). Ainsi, pour une durée fixe de la voyelle d'une syllabe VC, les jugements passent de 100 pc /juz/ à 70 pc /jus/ quand la durée de la friction passe de 50 à 250 ms.

h) L'intensité relative du bruit est aussi un facteur de la surdité. Dans ¹⁶, les syllabes synthétiques faites de la combinaison d'une explosion et d'une voyelle sans transitions étaient entendues sourdes. Dans ⁵², la présence d'explosions dans des syllabes qu'on voulait rendre sonores pour des oreilles américaines a obligé à compenser par l'addition d'une fondamentale, suivie d'une lenteur relative de T1. Dans ²⁷, les variations d'intensité du bruit de friction avaient une influence sur la perception de la sonorité, mais elle était très faible.

i) L'influence de la durée de la voyelle relativement à la durée de la friction consonantique suivante est le sujet principal de ²⁷. La technique par synthèse permet de combiner plusieurs durées de voyelles avec plusieurs durées de frictions, tout en gardant fixes les facteurs de transitions. Les résultats sont clairs (mais ils ne s'appliquent qu'aux fricatives): plus la durée relative de la voyelle est longue, plus la consonne est perçue sonore. Ainsi, pour une certaine durée fixe de la friction, les jugements passent de 100 pc /jus/ à 65 pc /juz/ quand la voyelle passe de 50 à 200 ms.

j) Notons enfin que dans une étude où les consonnes sont identifiées après distortion par filtrage et par addition de bruit ²⁴, les distinctions «sourde-sonore» et «orale-nasale» résistent beaucoup mieux que les distinctions de lieux d'articulation.

Les Voyelles Orales

L'étude systématique des voyelles orales par la synthèse a fait l'objet de deux articles: ^{7,15}. Dans ⁷, on trouve les fréquences de formant nécessaires pour synthétiser les 16 principales voyelles cardinales au moyen de deux formants seulement. Ces 16 voyelles ont été choisies à l'oreille parmi 235 combinaisons appropriées de variations de F2 pour F1 fixe et de F1 pour F2 fixe. Sur le diagramme des voyelles qui se forme en portant les fréquences de F1 et F2 en abscisse et en ordonnée, il est curieux de remarquer que [æ] est en ligne avec [i e ε], tandis que [a] est en ligne avec [y ø œ].

Dans ¹⁵, on établit les faits fondamentaux suivants:

a) Dans la synthèse, deux formants suffisent à bien caractériser le timbre des voyelles, même des voyelles nasales.

b) Mais les voyelles *humaines* sont souvent identifiées par trois formants. Autrement dit, dans la parole humaine, F3 joue un rôle dans l'identification de certaines voyelles, voire toutes les voyelles qui ont un F2 de fréquence assez haute, c'est-à-dire qui ont F2 et F3 assez rapprochés. Ce sont en général les voyelles antérieures.

c) Dans la perception, il y a équivalence relative entre deux formants rapprochés et un seul formant à une fréquence moyenne entre les deux. Ainsi les voyelles postérieures sont identifiables au moyen d'un seul formant dont la fréquence est intermédiaire entre F1 et F2 (F3 est très faible pour les voyelles postérieures, et ne contribue guère qu'à leur «naturel», étant aussi très loin de F2, en fréquence). De même lorsque F2 et F3 sont rapprochés, comme dans les voyelles antérieures, la perception de leur somme équivaut à peu près à la perception d'un seul formant dont la fréquence serait intermédiaire entre F2 et F3.

d) Dans les voyelles *synthétiques* «antérieures» à deux formants, la fréquence de F2 est intermédiaire entre les fréquences de F2 et F3 des voyelles naturelles de même timbre. Ainsi, le timbre du [i] naturel dont les formants sont à 250, 2500 et 3000 cps se synthétise bien avec deux formants à 250 et 2750 cps environ. (Mais il va sans dire qu'il se synthétise encore mieux avec trois formants aux mêmes fréquences que la voyelle naturelle!)

e) Les formants en dessus de 3000 cps pour [i] et en dessus de 2500 cps pour les autres voyelles n'ont guère de rôle dans la caractérisation linguistique des voyelles. Ils contribuent surtout à la caractérisation du timbre de la voix individuelle.

f) Les variations individuelles d'intensité des formants ont deux effets différents selon que les formants sont rapprochés ou éloignés (en fréquence). Si les deux formants dont on varie les intensités relatives sont éloignés l'un de l'autre, à mesure que la différence d'intensité s'accroît le timbre vocalique devient plus vague, puis il perd son identité linguistique pour prendre une identité musicale (généralement une dissonance de sons contigus); si les deux formants sont proches l'un de l'autre (en fréquence), la voyelle change de timbre comme si l'effet perceptuel de la somme des deux formants devenait de plus en plus semblable à l'effet perceptuel du formant qui avait gardé son intensité originelle.

g) Quand F1 seul diminue d'intensité, le changement de timbre

est perçu comme allant vers la nasalité. (La première indication du fait que l'indice acoustique de la nasalité vocalique est dans la faible intensité de F1 se trouve donc dans ¹⁵.)

L'étude ³⁸ des voyelles américaines par filtrage des fréquences qui sont supérieures à 670 cps (dans le but d'omettre tous les formants en dessus de F1) conclut que les indices acoustiques sont F1, F2, et la durée (deux degrés de durée). (Le rôle de la durée semble en effet indispensable pour distinguer deux voyelles comme /e/ et /I/ qui ont presque les mêmes fréquences de formants. Notons pourtant que le rôle de F3 n'a pas été inclus dans cette étude.)

Les fréquences des formants vocaliques ont été étudiés *par l'analyse* pour nombre de langues. Par exemple, pour les voyelles américaines, dans ^{1, 3, 14} et surtout ¹⁰; pour les voyelles françaises, dans ² et ⁸; pour les voyelles danoises, dans ²³; pour les voyelles suédoises, dans ³⁵; pour les voyelles polonaises, dans ⁴³; pour les voyelles japonaises, dans ⁴⁴.

Les analogues électriques décrits dans ^{5, 25, 29} ont produit de bonnes voyelles synthétiques, et ont ainsi contribué, non seulement à spécifier les corrélations articulatoires et acoustiques, mais à vérifier la valeur linguistique des résultats obtenus par la synthèse.

Enfin, une théorie proposée dans ³ vient d'être confirmée dans ⁴⁷ par la synthèse. L'identification linguistique des voyelles ne dépendrait pas entièrement de la fréquence *absolue* des formants, mais de leur fréquence relativement à la structure totale des formants du sujet parlant, structure qui peut légèrement varier d'une personne à l'autre, comme l'indiquent les divergences entre hommes, femmes, et enfants (fréquences légèrement plus hautes, dans cet ordre) établies dans ¹⁰.

Les Voyelles Nasales

Les indices de la nasalité vocalique ont été découverts grâce à la technique de synthèse ¹⁵ et surtout ²⁰, et confirmés plus tard par un analogue électrique de la bouche ⁴⁰, et par une analyse des voyelles nasales japonaises ⁴¹.

Le premier indice, le seul qui soit capable de transformer une voyelle orale en une voyelle nasale, indépendamment des autres indices, c'est la réduction d'intensité de F1. Pour les voyelles nasales françaises, par synthèse, il faut une réduction de 12 à 15 db relativement à l'intensité normale de F1 dans les voyelles orales.

Le second indice (second en importance) est un formant à environ 250 cps, que nous appellerons FN1 (premier formant nasal). C'est vraisemblablement le formant qui tient la première place dans la tenue des *consonnes* nasales. On sait qu'il contribue considérablement à la nasalité vocalique parce que lorsque FN1 est présent, il faut moins de réduction d'intensité de F1 pour que la voyelle soit identifiée comme nasale. Mais par lui-même il ne nasalise que très légèrement les voyelles.

Les autres indices, pas toujours visibles sur les spectrogrammes, sont très faibles et leurs effets perceptuels presque négligeables : ce sont principalement un formant vers 1000 cps et un autre vers 2000 cps.

On attribue hypothétiquement la réduction d'intensité de F1, soit au grand amortissement des cavités fibreuses du nez, qui agirait seulement sur les ondes de fréquence basse au niveau de F1 ⁴⁰; soit à des antirésonances qui supprimeraient une portion des tons de F1 ⁴¹. Les cavités nasales ayant un volume assez fixe, pour que ces antirésonances coïncident avec les fréquences de F1, il faudrait que les cavités buccales, surtout la cavité pharyngienne, fassent une accommodation de volume qui accorde les fréquences de F1 à celles des antirésonances.

L'hypothèse de l'amortissement est soutenue par le fait que les efforts de l'analogue pour produire des voyelles nasales ont d'abord échoué. La simple addition d'une troisième cavité, ne faisait que produire un formant additionnel vers 1000 cps, et les voyelles n'en étaient pas perceptuellement nasalisées. Pour arriver à produire des voyelles entendues comme nasales et ayant sur le spectre un F1 de très basse intensité, il a fallu donner un grand amortissement à la cavité nasale.

D'autre part, l'hypothèse des antirésonances, qui demanderait une accommodation des cavités, est soutenu par le fait que le F1 faible tend, pour toutes les voyelles nasales, vers une *même* fréquence d'environ 500 cps ⁴¹. Cela expliquerait l'évolution de toutes les voyelles nasales vers un *même* degré d'ouverture (mi-ouvertes) : au cours de l'histoire du français, [in yn un] deviennent [ẽ œ ÿ] (approximativement) et [ã] devient une voyelle qui se rapproche plus de [õ] que de [ã].

Prosodie

Les éléments prosodiques de la parole, tels que l'accent, le rythme, et l'intonation, commencent à être étudiés par la synthèse.

On en attend des résultats importants car les facteurs objectifs de durée, d'intensité, et de fréquence peuvent être variés de façon indépendante, et les résultats des variations isolées et combinées peuvent se juger subjectivement à l'oreille.

Une première étude a voulu comparer les effets des variations de durée et d'intensité (sans inclure, à ce point, la fréquence) dans la perception de la place de l'accent anglais. On a employé des mots tels que «object» qui sont compris comme substantif quand l'accent est sur la première syllabe et comme verbe quand il est sur la seconde. Quand les sujets identifiaient le mot comme substantif, on pouvait admettre qu'ils avaient perçu l'accent sur la première syllabe, et inversement. Résultats: les deux facteurs contribuent à la perception de la place de l'accent, mais la durée plus que l'intensité²⁸.

Des études comprenant, de plus, le facteur de la fréquence sont en cours.

Bibliographie Générale

Notons, pour terminer, les travaux généraux qui ont récapitulé et fait des hypothèses se rapportant aux indices acoustiques de la parole, entre 1947 et 1957. Dans l'ordre chronologique, c'est d'abord¹, avec une richesse de spectrogrammes de grande valeur, bien qu'ils correspondent en général à de la parole ralentie. Ensuite³, dont la perspicacité a mis sur la voie des recherches actuelles. Puis⁹, qui a proposé des traits distinctifs à caractère binaire en se basant en grande partie sur les indices acoustiques qui leur correspondent. Ce travail «préliminaire» s'étant fié à une analyse trop hâtive des spectres, sera entièrement à refaire quand les chercheurs acoustiques et physiologiques auront déterminé les vrais indices. Dans^{11, 12, et 13}, on peut suivre le développement rapide des recherches par la synthèse, mais les hypothèses, basées sur des résultats partiels, sont en partie fautives, car elles précèdent la découverte du concept du Locus de F2 et des transitions de F3. Enfin, dans⁴⁶ les spéculations prennent une forme plus avancée. On y trouve surtout les vues auxquelles ont mené ces dix années de travaux sur les rôles respectifs du niveau acoustique et du niveau articulaire dans la perception de la parole: en particulier, l'onde acoustique ne serait pas perçue directement, mais indirectement par référence au geste articulaire.

Conclusion

Bien que les progrès des dix années passées soient impressionnants, on est loin de pouvoir encore faire un tableau sûr et complet des indices acoustiques de la parole. Il reste non seulement à étudier à fond nombre de facteurs à peine explorés, mais à compléter le travail pour ceux qui sont les mieux connus en les étudiant maintenant en toutes positions et devant toutes les voyelles appropriées. La recherche portera sur plusieurs indices à la fois dans chacun des divers laboratoires mais elle sera quand même lente: dans l'étude complète d'un seul indice, il se passe généralement plusieurs années entre le moment où il est isolé et celui où les tests définitifs sont analysés. Nous ferons un autre rapport dès qu'une avance significative aura été réalisée.

Bibliographie

1. *Potter, R. K.; Kopp, G. A. and Green, H. C.*: Visible Speech (Van Nostrand, New York 1947).
2. *Delattre, P.*: Un triangle acoustique des voyelles orales du français. *French Rev.* 21: 477-484 (1948).
3. *Joos, M.*: Acoustic Phonetics (Waverly Press, Baltimore 1948).
4. *Cooper, F.*: Spectrum Analysis. *J. Acoust. Soc. Amer.* 22: 761-762 (1950).
5. *Dunn, H.*: Calculation of Vowel Resonances, and an Electrical Vocal Tract. *JAS.* 22: 740-753 (1950).
6. *Cooper, F. S.; Liberman, A. M. and Borst, J. M.*: The Interconversion of Audible and Visible Patterns as a Basis for Research in the Perception of Speech. *Proc. nat. Acad. Sci., Wash.* 37: 318-325 (1951).
7. *Delattre, P.; Liberman, A. M. and Cooper, F. S.*: Voyelles synthétiques à deux formants et voyelles cardinales. *Maître Phonét.* 96: 30-37 (1951).
8. *Delattre, P.*: The Physiological Interpretation of Sound Spectrograms. *Publ. Mod. Lang. Assoc. Amer.* 66: 864-876 (1951).
9. *Jacobson, R.; Fant, C. and Halle, M.*: Preliminaries to Speech Analysis, the Distinctive Features and their Correlates (Acoustics Laboratories of MIT, Cambridge, Mass. 1952).
10. *Peterson, G. and Barney, H.*: Control Methods Used in a Study of the Vowels. *JAS.* 24: 175-185 (1952).
11. *Delattre, P.; Liberman, A. M., Cooper, F. S. and Gerstman, L.*: Speech Synthesis as a Research Technique. *Proc. 7th Int. Congr. Ling. London 1952*; pp. 555-561 (1952).
12. *Delattre, P.; Cooper, F. S. and Liberman, A. M.*: Some Suggestions for Language Teaching Methods Arising from Research on the Acoustic Analysis and Synthesis of Speech. *Rep. 3rd. ann. Round Table Meet. Linguist. Lang. Teach.* 2: 31-47 (1952).
13. *Cooper, F. S.; Delattre, P., Liberman, A. M., Borst, J. M. and Gerstman, L.*: Some Experiments on the Perception of Synthetic Speech Sounds. *JAS.* 24: 597-606 (1952).
14. *Peterson, G.*: Information-Bearing Elements of Speech. *JAS.* 24: 629-636 (1952).

15. *Delattre, P.; Liberman, A. M., Cooper, F. S. and Gerstman, L.*: An Experimental Study of the Acoustic Determinants of Vowel Color; Observations on One- and Two-Formant Vowels Synthesized from Spectrographic Patterns. *Word* 8: 195-211 (1952).
16. *Liberman, A. M.; Delattre, P. and Cooper, F. S.*: The Role of Selected Stimulus-Variables in the Perception of the Unvoiced Stop Consonants. *Amer. J. Psychol.* 65: 497-517 (1952).
17. *Durand, M.*: De la formation des voyelles nasales. *Studia Linguist.* 7: 33-53 (1953).
18. *Cooper, F. S.*: Some Instrumental Aids to Research on Speech. Report of the Fourth Annual Round Table MLLT 3: 46-54 (1953).
19. *Schatz, C.*: The Role of Context in the Perception of Stops. *Language* 30: 47-57 (1954).
20. *Delattre, P.*: Les attributs acoustiques de la nasalité vocalique et consonantique. *Studia Linguist.* 8: 103-109 (1954).
21. *Liberman, A. M.; Delattre, P., Cooper, F. S. and Gerstman, L.*: The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants. *Psychol. Monogr.* 379: 1-14 (1954).
22. *Durand, M.*: La perception des consonnes occlusives: problèmes de palatalisation et de changements consonantiques. *Studia Linguist.* 8: 110-123 (1954).
23. *Fischer-Jørgensen, E.*: Acoustic Analysis of Stop Consonants. *Misc. Phonet.* 2: 42-59 (1954).
24. *Miller, G. and Nicely, P.*: Analysis of Perceptual Confusions Among some English Consonants. *JAS.* 27: 338-353 (1955).
25. *Stevens, K. and House, A.*: Development of a Quantitative Description of Vowel Articulation. *JAS.* 27: 484-494 (1955).
26. *Delattre, P.; Liberman, A. M. and Cooper, F. S.*: Acoustic Loci and Transitional Cues for Consonants. *JAS.* 27: 769-774 (1955).
27. *Denes, P.*: Effect of Duration on the Perception of Voicing. *JAS.* 27: 761-764 (1955).
28. *Fry, D.*: Duration and Intensity as Physical Correlates of Linguistic Stress. *JAS.* 27: 765-768 (1955).
29. *House, A. and Stevens, K.*: Auditory Testing of a Simplified Description of Vowel Articulation. *JAS.* 27: 882-887 (1955).
30. *Malmberg, B.*: The Phonetic Basis for Syllable Division. *Studia Linguist.* 9: 80-87 (1955).
31. *Hughes, G. and Halle, M.*: Spectral Properties of Fricative Consonants. *JAS.* 28: 303-310 (1956).
32. *Durand, M.*: De la perception des consonnes occlusives, questions de sonorité. *World* 12: 15-34 (1956).
33. *Liberman, A. M.; Delattre, P., Gerstman, L. and Cooper, F. S.*: Tempo of Frequency Change as a Cue for Distinguishing Classes of Speech Sounds. *J. exp. Psychol.* 52: 127-138 (1956).
34. *Borst, J.*: The Use of Spectrograms for Speech Analysis and Synthesis. *J. Audio Engng. Soc.* 4: 14-23 (1956).
35. *Malmberg, B.*: Distinctive Features of Swedish Vowels; Some Instrumental and Structural Data. For R. Jacobson, pp. 316-321 (1956).
36. *Fischer-Jørgensen, E.*: The Commutation Test and its Application to Phonemic Analysis. For R. Jacobson, pp. 140-151 (1956).
37. *Householder, F.*: Unreleased ptk in American English. For R. Jacobson, pp. 235-244 (1956).
38. *Miller, G.*: The Perception of Speech. For R. Jacobson, pp. 353-360 (1956).
39. *Malécot, A.*: Acoustic Cues for Nasal Consonants; an Experimental Study Involving a Tape-Splicing Technique. *Language* 32: 274-284 (1956).

40. House, A. and Stevens, K.: Analog Studies of the Nasalization of Vowels. J. Speech Dis. 21: 218-232 (1956).
41. Hattori, S.; Yamamoto, K. and Fujimura, O.: Nasalization of Vowels and Nasals. Rep. Kobayashi sci. Inst. 6: 226-235 (1956).
42. Stevens, K. and House, A.: Studies of Formant Transitions Using a Vocal-Tract Analog. JAS. 28: 578-585 (1956).
43. Jassem, W.: The Formants of Sustained Polish Vowels; I Preliminary Study. Study of Sounds; pp. 335-349 (Chiyoda, Tokio 1957).
44. Hattori, S.; Yamamoto, K., Kohasi, Y and Fujimura, O.: Vowels of Japanese. Rep. Kobayashi sci. Inst. 7: 69-79 (1957).
45. Halle, M.; Hughes, G. and Radley, J.-P.: Acoustic Properties of Stop Consonants. JAS. 29: 107-116 (1957).
46. Liberman, A. M.: Some Results of Research on Speech Perception. JAS. 29: 117-123 (1957).
47. Ladefoged, P. and Broadbent, D.: Information Conveyed by Vowels. JAS. 29: 98-104 (1957).
48. Lisker, L.: Closure Duration and the Intervocalic Voiced-Voiceless Distinction in English. Language 33: 42-49 (1957).
49. O'Connor, J. D.; Gerstman, L., Liberman, A. M., Delattre, P. and Cooper, F. S.: Acoustic Cues for the Perception of Initial /wɹl/ in English. Word 13: 24-44 (1957).
50. House, H.: Analog Studies of Nasal Consonants. J. Speech Dis. 22: 190-204 (1957).
51. Gerstman, L.: Cues for Distinguishing among Fricatives, Affricate, and Stop Consonants. Diss. (New York University 1957) (Research done at Haskins Laboratories, New York).
52. Hoffmann, H.: A Study of Some Cues in the Perception of the Voiced Stop Consonants. Diss. (University of Connecticut 1957) (Research done at Haskins Laboratories, New York).

Adresse de l'auteur: Pierre Delattre, 515 15th Street, Boulder, Colorado (USA)

Libri

Franklin Brook: Stammering and its Treatment. Pitman Medical Publishing Co. Ltd., London 1957, 134 p. 21/-.

It is difficult for a physician to criticize a book on stammering that is devoid of any exact symptomology of the speech impediment. Medical science is based upon thorough description of symptoms and signs. This is fundamental to all the other means by which medicine tries to understand, to treat and to cure any disease. Numerous publications on stammering have appeared, especially more recently, which are wanting in this respect, among them this one of *Franklin Brook*. Therefore, only with some restrictions can a review be offered which pays tribute to the work, the enthusiasm and the series of recommendations and warnings of the writer.

In reference to *Brook's* quotation of *Johnson's* statement that no stammerer has been found among the North American Indians the reviewer consulted an expert on Indian language some time ago and received a list of ten Indian words that mean stammering! "Stammering is often found to be a symptom of emotional disturbance", but "the emotional background of stammering is very confused". True, because stammering will

inevitably bring the sufferer into most painful situations – so starting a vicious circle. To prove the influence of factors other than psychogenic, one factor which the author mentions is that feverish illness causing disturbances of the central nervous system may be responsible. Here the discussion between *H. Gutzmann* and this reviewer about stammering as focal symptom should have been taken into consideration. The statement that in many cases stammering was present “ever since the child started talking” is, according to the experience of the reviewer, not correct but due to an inexact anamnesis. The repetitions of initial syllables in children from 12 to 27 months are not called, as *Brook* believes, “clonic” but “primary cloni”. The question of a combination of stammering and dyslalia is dealt with as well as many other important questions. It should have been mentioned that a severe dyslalia must be treated together with the stammering and that this necessity might produce complications. “Having recognized, for instance, that the letter k is a boggy letter, he will proceed to stammer whenever he meets the letter.” This statement should have been completed by emphatically pointing to the fact that stammering never occurs at the end of a word (unless in the one case of *van Thal* who used sometimes a final syllable as a starter for a following word). From this fact quite important conclusions can, in the reviewer’s opinion, be drawn concerning the “core of stuttering”. Much has been written in this book and others about the difference about the older and modern methods of treatment, the first being only a speech therapy, the latter a complex one which “demands a lot of hard work spread over a long period of at least 12 months”. Innumerable stammerers have been cured by the old methods within a fraction of this time! The treatment which *Brook* uses and recommends includes: General relaxation (this, as some other parts of the book, must be read in the original). “When a group of muscles is working at minimum tension and the remaining body musculature is relaxed as far as possible, we are in a state of *differential relaxation* and can continue a movement for a long period without strain or fatigue. In practising differential relaxation therefore, the stammerer will be seeking the greatest amount of ease in speaking.” The term “*the stammerer*” is frequently used in this book when, in the reviewer’s opinion, “*some stammerers*” should be used. “If the musculature is sufficiently relaxed, ... there will be a fair degree of clarity loss on all consonants and a slight clarity loss on the vowels.” Should not the “*Ventriloquism*” method have been quoted here? On page 68 only a clue is given to the development of signs. A much more exact description resulted from many years of work of the former Viennese school. *Voluntary stammering* plays a big part in *Brook’s* method. An exact knowledge of the so-called *imitation stammering* which differs essentially in its signs from other forms of stammering could contribute much to a better understanding of the influence of voluntary stammering upon the patient. *Voluntary Stammering Techniques* is the title of one chapter. Also recommended is the Stop Go method consisting of stopping whenever the patient feels like stopping. The writer states that group treatment is often helpful. A list of suggestions is offered on page 102. Among other very laudable attempts of the writer is the one to deal thoroughly with the environment of stammerers, especially home and school, and to show how they can help the therapist and in this way the patient. Thorough reading of the book is necessary to fully appreciate its merits and its shortcomings. Despite the latter, the volume will be a great help to some therapists. *Brook* points to the fact that the choice of treatment used will depend to a great extent on the whole personality of the therapist. Scientific findings should certainly influence that choice! Therefore, a knowledge of the literature is presupposed! *E. Froeschels*, New York

Etudes Tsiganes: *Bulletin de l’Association des Etudes Tsiganes* (5, rue Las-Cases, Paris VII^e).

Diese neue Zeitschrift erscheint zwar bereits seit dem 15. April 1955, ist aber außerhalb Frankreichs bisher kaum bekannt geworden. Entgegen dem ursprünglichen Plan

sind bis zum Januar 1957 auch nur 8 Hefte herausgekommen. Sie sind von wechselndem Umfang und inhaltlich von ungleichem Wert. Offensichtlich fällt es der Vereinigung schwer, ihr Periodikum gleichberechtigt gegen das internationalen und traditionellen Ruf genießende «Journal of the Gypsy Lore Society» zu behaupten. Trotz allen noch offenbleibenden Wünschen für die redaktionelle Gestaltung wäre es jedoch bedauerlich, die Etudes Tsiganes etwa wieder verschwinden zu sehen. Beim Mangel an neuer, ernst zu nehmender Literatur über die Zigeuner – an belletristischen Buchreportagen ist derzeit eine unerwünschte Hochflut zu verzeichnen – ist jede gehaltvolle Zeitschrift willkommen. Frankreich könnte durch die Vielzahl seiner Zigeunerstämme bei sorgfältiger, sachgemäßer Forschung Entscheidendes zur Lösung folkloristischer und linguistischer Probleme beitragen. Es scheint dem jungen Periodikum aber an philologischen Mitarbeitern zu mangeln. Abgesehen von den summarisch und korrekt informierenden «Observations sur la langue tsigane» von *P. Meille*, die nur dem Laien Neues bringen, verdient nur eine Erzählung des in seinem Kalderaš-Dialekt schreibenden *M. Maximoff*, «E pasledno vudar» (Die letzte Tür), Erwähnung. Es steht zu hoffen, daß sich die Vereinigung baldmöglichst entscheidet, ob sie ihrem Organ durch gute linguistische Originalbeiträge die Lebensfähigkeit und das Interesse der internationalen Zigeunerphilologie sichern will.

S. A. Wolf, Berlin

Zdeňka Jamnická-Šmerglová: Dějiny našich cikánů. Orbis, Praha 1955.

In der Tschechoslowakei standen die Behörden 1945 vor dem Problem, etwa 150 000 nomadisierende Zigeuner an Selbsthaftigkeit und regelmäßige Arbeit zu gewöhnen. Die Verfasserin nimmt das – in der «Geschichte unserer Zigeuner» – zum Anlaß, die traurige Vergangenheit und die glückliche Gegenwart der Zigeuner gegenüberzustellen. Als roter Faden durchzieht alle Ausführungen die These, daß nur im Gesellschaftsaufbau volksdemokratischer Staaten die endgültige und befriedigende Lösung des Zigeunerproblems möglich ist. Der gleichen Ansicht ist z. B. auch der rühmlichst bekannte Zigeunerphilologe und -folklorist Prof. *Uhlik* in Sarajevo. Aber *Uhlik* ist sicherlich der beste Zigeunerkenner Südosteuropas und wird seine Meinung subjektiv und objektiv verteidigen können, wenn es gewünscht wird. Die *Jamnická-Šmerglová* bringt nichts Eigenes außer einer Fülle von Zusammengestopfeltem aus älteren Autoren, Entstellungen und Verfälschungen sowie als Tatsachen präsentierten subjektiven Vermutungen. So «weiß» sie, daß die Zigeuner, «die Avantgarde des indischen Proletariats», gen Westen wanderten, um sich der «sozialen Unterdrückung» zu entziehen! Diese und ähnliche Phrasen hat der ernsthafte Zigeunerforscher Dr. *Lípa*, der selber durchaus «die wirkliche marxistische Analyse der Tatsachen» bejaht, bereits in der offiziellen «Československá ethnografie» als einen «Mißbrauch der marxistischen Terminologie» bezeichnet.

Die philologische Unkenntnis der Verfasserin übertrifft womöglich noch ihre historischen Mängel. Die klassischen westeuropäischen Zigeunerphilologen übergeht sie mit Schweigen. Da sie jedoch von *Sowa* und seine bekannte Untersuchung des slowakischen Zigeunerdialekts nicht ignorieren kann, macht sie ihn aus einem Deutschen zum Polen. Die «walachischen Zigeuner», von denen *Gjorgjevič* schon 1903 schrieb, daß sie nur das Rumänische beherrschten, läßt die *Jamnická-Šmerglová* 1955 plötzlich wieder die Zigeunersprache sprechen, wenn auch wenigstens «von rumänischen Sprachelementen durchsetzt». Für die von ihr konstruierten Gruppen der «schwarzen» und «weißen» Zigeuner hat sie für die letztere auch gleich eine angebliche Gaunersprache mitgeliefert, deren einzige Grundlage in Unorientiertheit über die tschechische Gaunersprache besteht.

Man muß sich dem tschechischen Urteil anschließen, daß Druck und Erscheinen dieses Buches ein «Irrtum» waren.

S. A. Wolf, Berlin

Bohuslav Hála: Nature acoustique de voyelles. Acta universitatis Carolinae, Prag 1956.

Die tschechischen Vokale werden wie folgt charakterisiert:

Zwischen den geschlossenen und den offenen Vokalen nehmen sie eine Zwischenstellung ein. Das tschechische *e* befindet sich zwischen dem geschlossenen *e* (fr. *été*, deutsch *See*, engl. *hay*) und dem offenen *e* (fr. *tête*, deutsch *Bär*, engl. *there*). Das *o* ist weder geschlossen wie fr. *beau*, deutsch *Sohn*, engl. *know* noch offen wie fr. *bord*, engl. *door*. Die langen Laute sind leicht etwas geschlossener als die kurzen. Die tschechischen *a*-Laute befinden sich zwischen dem vorderen *a* (fr. *madame*, engl. *man*) und dem hinteren (fr. *pas*, engl. *father*, deutsch *Vater*). Das lange *a* ist um eine Nuance mehr hinten, das kurze *a* mehr nach vorn gesprochen.

Die Spektralanalyse der tschechischen Vokale hat folgendes ergeben:

1. Die Formantfrequenzen erstrecken sich von 350 bis 2230 Hz ($2\frac{1}{2}$ Oktaven), im Gegensatz zu den französischen Vokalen mit 4 Oktaven Umfang.

2. Die Formanten sowie höhere Frequenzen sind wie folgt verteilt:

Formant 1 Mund-Res.	Formant 2 Rachen-Res.	Zwischen- resonanz	Ober-Res.	Hochton-Res.
ú 350	680		2230	3155
u 385	758		2100	2889
ó 530	900		1792	2734
o 580	982		1617	2610
á 1175	795	987	1577	2288
a 1280	750	1022	1776	2369
e 1660	572		2070	2563
é 1750	510		2120	2700
i 2120	355		2365	2833
í 2230	326		2610	3400

ú langer Vokal, u kurzer Vokal

Die Ergebnisse stimmen in großen Zügen mit früheren Messungen anderer Autoren überein, die hier aufgeführt wurden.

Vor Beginn dieser Messung hat der Verfasser 16 ältere Methoden der Formantbestimmung überprüft: 1. Flüstermethode, 2. Erregung der Mundhöhle durch Druckluft, 3–6. verschiedene Resonatoren nach *Rousselot* und *Schäfer*, 7–8. Tonometer nach *König* und *Schäfer*, 9. photographische Registrierung, 10–13. verschiedene Registrierung mittels Tonfilm, 14. Kathodenstrahl-Oszillograph, 15. Strobilion nach *Scripture*, 16. Geschwindigkeitsänderungen der Schallplattenwiedergabe.

Die moderneren Methoden der Filter- und Suchtonanalyse wie auch *VisibleSpeech* wurden nicht berücksichtigt. Die Betrachtungen über die Stabilität und die Entstehung der Formanten stützen sich auf die älteren klassischen Anschauungen. Es wird noch einmal auf die Auseinandersetzungen *Helmholtz–Hermann* eingegangen und die bekannte Lösung bestätigt, wonach beide Anschauungen durch die Beachtung eines Formantbereichs zu ihrem Recht kommen.

F. Winckel, Berlin

Directives sur la

Rédaction d'un Manuscrit et l'établissement d'une Bibliographie

avec une Liste des

Abréviations des Périodiques médicaux les plus courants

SOMMAIRE

A. Manuscrit	1
1. Généralités	1
2. Articles de périodiques	1
3. Manuscrit d'un livre	2
B. Illustrations	2
C. Correction des épreuves	3
Schéma	3
D. Bibliographie	5
1. Travaux publiés dans des périodiques	5
2. Livres, thèses, monographies, contributions à des manuels et précis, rapports de congrès et symposiums, etc.	6
3. Exemples	7
E. Abréviations	8
1. Règles générales	8
2. Abréviations générales	8
3. Liste des titres abrégés	9

Vous pouvez recevoir les directives en nous faisant parvenir 4 coupons-réponse
internationaux ou des timbres-poste d'une valeur de fr.s. 2.—

Eine deutsche Ausgabe ist bereits erschienen

English edition in preparation

S. KARGER Arnold-Böcklin-Strasse 25 BASEL 11

Journals - Zeitschriften - Revues

Acta Anatomica (Int. Archiv für Anatomie, Histologie, Embryologie und Zellforschung – Int. Archives of Anatomy, Histology, Embryology and Cytology – Archives Int. d'Anatomie, d'Histologie, d'Embryologie et de Cytologie)

Acta Genetica et Statistica Medica

Acta Haematologica (Int. Zeitschrift für Hämatologie – Int. Journal of Haematology – Journal Int. d'Hématologie)

Acta Psychotherapeutica Psychosomatica et Orthopaedagogica (Int. Journal of Psychotherapy, Psychosomatics, Special Education – Int. Zeitschrift für Psychotherapie, Psychosomatik und Heilpädagogik – Journal Int. de Psychothérapie, Psychosomatique, Education Spéciale)

Annales Paediatrici (Jahrbuch für Kinderheilkunde – Int. Review of Pediatrics – Revue Int. de Pédiatrie)

Cardiologia (Int. Archiv für Kreislaufforschung – Int. Archives of Cardiology – Archives Int. du Cœur et des Vaisseaux)

Confinia Neurologica (Grenzgebiete der Neurologie – Borderland of Neurology – Les Confins de la Neurologie)

Confinia Psychiatrica (Grenzgebiete der Psychiatrie – Borderland of Psychiatry – Les Confins de la Psychiatrie)

Dermatologica (Dermatologische Zeitschrift – Int. Journal of Dermatology – Journal Int. de Dermatologie)

Folia Phoniatica (Int. Zeitschrift für Phoniatrie – Int. Journal of Phoniatriy – Journal Int. de Phoniatrie)

Gastroenterologia (Int. Zeitschrift für Gastroenterologie – Int. Review of Gastroenterology – Revue Int. de Gastroentérologie)

Gerontologia (Zeitschrift für experimentell-biologische und -medizinische Altersforschung – Journal of Experimental Biological and Medical Research on Ageing – Journal de Recherches expérimentales, biologiques et médicales sur le vieillissement)

Gynaecologia (Monatsschrift für Geburtshilfe und Gynäkologie – Int. Monthly Review of Obstetrics and Gynecology – Revue Int. Mensuelle d'Obstétrique et de Gynécologie)

International Archives of Allergy and Applied Immunology

Oncologia (Zeitschrift für Erforschung, Bekämpfung, Behandlung und Soziologie der Krebskrankheit – Journal of Cancer Research, Prevention, Treatment and Sociological Aspect – Revue de l'Exploration, du Traitement et de la Sociologie du Cancer)

Ophthalmologica (Int. Zeitschrift für Augenheilkunde – Int. Journal of Ophthalmology – Journal Int. d'Ophthalmologie)

Phonetica (Int. Zeitschrift für Phonetik – Int. Journal of Phonetics – Journal Int. de Phonétique)

Practica Oto-Rhino-Laryngologica (Int. Zeitschrift für Hals-, Nasen-, Ohrenheilkunde und ihre Grenzgebiete – Int. Review of Otolaryngology – Revue Int. d'Otolaryngologie)

Psychiatria et Neurologia (Internationale Monatsschrift für Psychiatrie und Neurologie – Revue Internationale Mensuelle de Psychiatrie et Neurologie – International Monthly Review of Psychiatry and Neurology)

Radiologia Clinica (Int. Radiologische Rundschau – Int. Radiological Review – Revue Int. de Radiologie)

Schweizerische Zeitschrift für Allgemeine Pathologie und Bakteriologie (Revue Suisse de Pathologie générale et de Bactériologie – Swiss Journal of General Pathology and Bacteriology)

Schweizerische Zeitschrift für Tuberkulose und Pneumonologie (Revue Suisse de la Tuberculose et de la Pneumologie – Rivista Svizzera della Tubercolosi e della Pneumologia)

Urologia Internationalis

Vita Humana (Int. Zeitschrift für Lebensaltersforschung – Int. Journal of Human Development – Journal Int. de Développement Humain)

Vox Sanguinis (Journal of Blood Transfusion and Immunohaematology – Journal de Transfusion Sanguine et d'Immunohématologie – Zeitschrift für Bluttransfusion und Immunohämatologie)



S. KARGER AG

BASEL / NEW YORK